OM protein - protein search, using sw model

Run on: March 28, 2005, 06:30:29 ; Search time 123 Seconds
(without alignments)
267.273 Million cell updates/sec

Title: US-10-071-645-2
Perfect score: 453
Sequence: 1 MWEVLPYGDEKLSPYGDGGD.....VLKNMTDKAPLVSNSPKTMS 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	453	100.0	85	5	AAE19431	Aae19431 Human PAP
2	391	86.3	78	6	ABU65229	Abu65229 Novel hum
3	259.5	57.3	79	6	ABR40376	Abr40376 Calmoduli
4	253.5	56.0	79	5	AAU77100	Aau77100 Human cal
5	133	29.4	27	6	ABR40377	Abr40377 Calmoduli
6	71.5	15.8	148	4	AAG83014	Aag83014 S. epider
7	71.5	15.8	217	5	ABP40348	Abp40348 Staphyloc
8	71.5	15.8	217	8	ADS04904	Ads04904 Staphyloc
9	71	15.7	145	2	AAR10720	Aar10720 (Gly59)IL
10	68.5	15.1	140	5	ABU50950	Abu50950 Helicobac
11	68.5	15.1	142	3	AAB52498	Aab52498 Helicobac
12	68.5	15.1	155	5	ABU51049	Abu51049 Helicobac
13	68.5	15.1	157	3	AAB52603	Aab52603 Helicobac
14	68.5	15.1	289	5	ABU52122	Abu52122 Helicobac
15	68.5	15.1	659	5	ABU51075	Abu51075 Helicobac
16	68.5	15.1	1943	6	ABU30826	Abu30826 Protein e
17	68.5	15.1	1974	2	AAW98391	Aaw98391 H. pylori
18	68	15.0	112	2	AAR69760	Aar69760 pMON5973
19	68	15.0	117	5	ABU52010	Abu52010 Helicobac
20	67.5	14.9	3194	6	ABU31122	Abu31122 Protein e
21	67	14.8	114	2	AAR69752	Aar69752 pMON5979
22	67	14.8	932	8	ADN21084	Adn21084 Bacterial
23	66.5	14.7	252	6	ABM68267	Abm68267 Photorhab
24	66	14.6	105	2	AAR69761	Aar69761 pMON5974
25	66	14.6	333	8	ADS10752	Ads10752 Human the
26	66	14.6	575	8	ADS23487	Ads23487 Bacterial
27	65	14.3	113	4	AAE13935	Aae13935 Myelopoie
28	65	14.3	133	2	AAR09326	Aar09326 Sequence
29	65	14.3	133	8	ADL89144	Adl89144 Human mod
30	65	14.3	133	8	ADL89151	Adl89151 Human mod
31	65	14.3	145	2	AAR10730	Aar10730 (Ala84)IL
32	65	14.3	145	2	AAR10731	Aar10731 (Ala16Ala
33	65	14.3	271	2	AAR26359	Aar26359 pIXY 321.
34	65	14.3	271	2	AAR20113	Aar20113 Human GM-
35	65	14.3	271	2	AAR23851	Aar23851 GM-CSF/IL
36	65	14.3	275	2	AAR11067	Aar11067 IL-3/GM-C
37	65	14.3	275	2	AAR26360	Aar26360 pIXY 344.
38	65	14.3	275	2	AAR20114	Aar20114 Human IL-
39	65	14.3	301	2	AAR37334	Aar37334 Human IL-
40	65	14.3	303	2	AAR37335	Aar37335 Human MGF
41	65	14.3	307	4	AAE14125	Aae14125 Myelopoie
42	65	14.3	322	4	AAE14126	Aae14126 Myelopoie
43	65	14.3	349	4	AAE14127	Aae14127 Myelopoie
44	64.5	14.2	248	7	ABO70749	Abo70749 Pseudomon
45	64	14.1	113	4	AAE13931	Aae13931 Myelopoie

ALIGNMENTS

RESULT 1
AAE19431
ID AAE19431 standard; protein; 85 AA.
XX
AC AAE19431;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human PAPAP protein.
XX
KW Human; PAPAP protein; schizophrenia candidate gene; g34872 gene;
KW schizophrenia; bipolar disorder; central nervous system disorder;
KW psychotic disorder; mood disorder; autism; mental retardation;
KW psychiatric disease; anxiety disorder; impulse-control disorder;
KW eating disorder; cognitive disorder; personality disorder; vaccine;
KW chromosome 1p35-p36; neuroleptic; antialcoholic; tranquiliser;
KW antidepressant; nootropic; antiaddictive.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Domain	8. .16
FT		/note= "ASP and GLU racemase motif"
FT	Region	8. .11
FT		/note= "Immunogenic epitope"
FT	Modified-site	18. .23
FT		/note= "N-myristoylation site"
FT	Modified-site	22. .27
FT		/note= "N-myristoylation site"
FT	Modified-site	26. .28
FT		/note= "Protein kinase C phosphorylation site"
FT	Region	31. .33
FT		/note= "Immunogenic epitope"
FT	Modified-site	37. .42
FT		/note= "N-myristoylation site"
FT	Region	40. .47
FT		/note= "Immunogenic epitope"
FT	Region	51. .54
FT		/note= "Immunogenic epitope"
FT	Modified-site	53. .55
FT		/note= "Protein kinase C phosphorylation site"
FT	Region	59. .62
FT		/note= "Immunogenic epitope"
FT	Modified-site	69. .72
FT		/note= "N-glycosylation site"
FT	Modified-site	71. .73
FT		/note= "Protein kinase C phosphorylation site"
FT	Region	80. .83

```
FT                  /note= "Immunogenic epitope"
FT   Modified-site  80. .82
FT                  /note= "Protein kinase C phosphorylation site"
XX
PN   WO200212279-A2.
XX
PD   14-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-IB001891.
XX
PR   07-AUG-2000; 2000US-0223482P.
XX
PA   (GEST ) GENSET.
XX
PI   Bihain B,  Bour B,  Bougueleret L;
XX
DR   WPI; 2002-241732/29.
DR   N-PSDB; AAD31052.
XX
PT   Novel isolated and purified or recombinant polynucleotide encoding PAPAP
PT   protein, useful for diagnosing and treating schizophrenia, bipolar
PT   disorder and other central nervous system disorders.
XX
PS   Claim 8; Page 93; 96pp; English.
XX
CC   The invention relates to human PAPAP polypeptides and polynucleotides.
CC   The invention also concerns the interaction of PAPAP with schizophrenia
CC   candidate gene g34872. PAPAP polypeptides, gene and anti-PAPAP antibodies
CC   are useful for treating schizophrenia, bipolar disorder or related
CC   central nervous system (CNS) disorders e.g. psychotic disorders, mood
CC   disorders, autism, substance dependence and alcoholism, mental
CC   retardation and other psychiatric diseases including cognitive, anxiety,
CC   eating, impulse-control and personality disorders. They are also used in
CC   vaccines. The present sequence is human PAPAP protein. The PAPAP gene is
CC   located on chromosome 1p35-p36
XX
SQ   Sequence 85 AA;

  Query Match             100.0%;  Score 453;  DB 5;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-51;
  Matches   85;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAPLVSNSPKTMS 85
              |||||||||||||||||||||||||
Db         61 DRIDDVLKNMTDKAPLVSNSPKTMS 85


RESULT 2
ABU65229
ID   ABU65229 standard; protein; 78 AA.
XX
AC   ABU65229;
XX
DT   16-MAY-2003  (first entry)
XX
DE   Novel human protein NOV9a.
XX
KW   Human; NOV; adrenoleukodystrophy; congenital adrenal hyperplasia;
KW   haemophilia; hypercoagulation; autoimmune disease; allergy;
KW   immunodeficiency; transplantation; Von Hippel-Lindau syndrome;
KW   Alzheimer's disease; stroke; tuberous sclerosis; hypercalcaemia;
KW   Parkinson's disease; Huntington's disease; cancer; fertility; diabetes;
KW   adult respiratory distress syndrome; infection; tissue typing;
KW   forensic identification.
XX
OS   Homo sapiens.
XX
PN   WO200290500-A2.
XX
PD   14-NOV-2002.
XX
PF   02-MAY-2002; 2002WO-US014256.
XX
PR   03-MAY-2001; 2001US-0288395P.
PR   07-MAY-2001; 2001US-0289087P.
PR   08-MAY-2001; 2001US-0289619P.
PR   09-MAY-2001; 2001US-0289817P.
PR   09-MAY-2001; 2001US-0289818P.
PR   11-MAY-2001; 2001US-0290194P.
PR   14-MAY-2001; 2001US-0290753P.
PR   15-MAY-2001; 2001US-0291189P.
PR   21-MAY-2001; 2001US-0292374P.
PR   23-MAY-2001; 2001US-0293107P.
PR   25-MAY-2001; 2001US-0293747P.
PR   29-MAY-2001; 2001US-0294110P.
PR   30-MAY-2001; 2001US-0294434P.
PR   10-SEP-2001; 2001US-0318346P.
PR   17-SEP-2001; 2001US-0322646P.
PR   01-MAY-2002; 2002US-00136728.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Spytek KA,  Li L,  Edinger SR,  Stone DJ,  Guo X,  Anderson DW;
PI   Patturajan M,  Gerlach VL,  Taupier RJ,  Pena CEA,  Padigaru M;
PI   Kekuda R,  Gorman L,  Zerhusen BD,  Smithson G,  Macdougall JR;
PI   Mezes PS,  Peyman JA,  Zhong M;
XX
DR   WPI; 2003-103511/09.
DR   N-PSDB; ABX97604.
XX
PT   New NOVX polypeptides and polynucleotides useful for treating or
PT   preventing e.g. congenital adrenal hyperplasia, hemophilia,
PT   hypercoagulation, autoimmune disease, allergies, immunodeficiencies,
PT   transplantation.
XX
PS   Claim 20; Page 107; 300pp; English.
XX
CC   The invention describes an isolated polypeptide, NOVX, comprising a
CC   sequence or a mature form of one of 21 51-1543 residue amino acid
CC   sequences (P1-P21), given in the specification. The NOVX polypeptides,
CC   polynucleotides and antibodies are useful in the manufacture of a
CC   medicament for treating or preventing e.g. adrenoleukodystrophy,
CC   congenital adrenal hyperplasia, haemophilia, hypercoagulation, autoimmune
CC   disease, allergies, immunodeficiencies, transplantation, Von Hippel-
CC   Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC   hypercalcaemia, Parkinson's disease, Huntington's disease, cancer,
CC   fertility, diabetes, adult respiratory distress syndrome, viral,
CC   bacterial and parasitic infections. The nucleic acid sequences may be
CC   used in chromosome mapping, identifying individual from minute biological
CC   samples (tissue typing), and in forensic identification of a biological
CC   sample. This is the amino acid sequence of a novel human protein (NOV)
XX
SQ   Sequence 78 AA;

  Query Match             86.3%;  Score 391;  DB 6;  Length 78;
  Best Local Similarity   98.7%;  Pred. No. 2.3e-43;
  Matches   74;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAP 75
              |||||||||||||||
Db         61 DRIDDVLKNMTDKAP 75


RESULT 3
ABR40376
ID   ABR40376 standard; protein; 79 AA.
XX
AC   ABR40376;
```

```
XX
DT   21-JUL-2003  (first entry)
XX
DE   Calmodulin kinase II inhibitor peptide AC3-I SEQ ID NO:4.
XX
KW   AC3-I; cardiomyopathy; myocardial contractile dysfunction;
KW   myocardial infarction; Calmodulin Kinase II; CaMKII; cardiant.
XX
OS   Synthetic.
XX
PN   WO2003029428-A2.
XX
PD   10-APR-2003.
XX
PF   01-OCT-2002; 2002WO-US031496.
XX
PR   01-OCT-2001; 2001US-0326576P.
PR   08-OCT-2001; 2001US-0328010P.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Anderson M;
XX
DR   WPI; 2003-371993/35.
XX
PT   Treating or preventing dilated cardiomyopathy or myocardial contractile
PT   dysfunction after myocardial infarction in a subject by administering an
PT   inhibitor of Calmodulin Kinase II.
XX
PS   Claim 4; Page 57; 58pp; English.
XX
CC   The invention relates to a novel method for treating or preventing
CC   dilated cardiomyopathy or myocardial contractile dysfunction after
CC   myocardial infarction in a subject. The method comprises administering an
CC   inhibitor of Calmodulin Kinase II (CaMKII). The method of the invention
CC   has cardiant activity. The method is useful for treating or preventing
CC   dilated cardiomyopathy or myocardial contractile dysfunction after
CC   myocardial infarction in a subject. The present sequence is used in the
CC   exemplification of the invention
XX
SQ   Sequence 79 AA;

  Query Match             57.3%;  Score 259.5;  DB 6;  Length 79;
  Best Local Similarity   65.4%;  Pred. No. 5.3e-26;
  Matches   51;  Conservative   10;  Mismatches   12;  Indels    5;  Gaps    2;

Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:||||::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYGEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 4
AAU77100
ID   AAU77100 standard; protein; 79 AA.
XX
AC   AAU77100;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human calmodulin-dependent protein kinase inhibitor.
XX
KW   Human; CaM-KIIN; calmodulin-dependent protein kinase inhibitor;
KW   treating stroke; head trauma; multiple sclerosis; Parkinson's disease;
KW   Alzheimer's disease; spinal cord injury; mental disorder; memory deficit;
KW   epilepsy; schizophrenia; bipolar disorder; cancer; ischaemia;
KW   lung disorder; malfunction of secretion; pancreatic disease; pain;
KW   immune system disorder; vaccine; gene therapy; chromosome 3;
KW   hereditary spastic paraplegia.
XX
OS   Homo sapiens.
XX
PN   WO200216415-A2.
XX
PD   28-FEB-2002.
XX
PF   23-AUG-2001; 2001WO-EP009748.
XX
PR   25-AUG-2000; 2000EP-00118541.
XX
PA   (MERE ) MERCK PATENT GMBH.
XX
PI   Duecker K;
XX
DR   WPI; 2002-280911/32.
DR   N-PSDB; ABK11646.
XX
PT   Novel calmodulin-dependent protein kinase inhibitor polypeptides useful
PT   for treating stroke, trauma, multiple sclerosis, Parkinson's disease,
PT   Alzheimer's disease, spinal cord injury, mental disorders, memory
PT   disorders.
XX
PS   Claim 1; Page 34; 34pp; English.
XX
CC   The invention relates to an isolated calmodulin-dependent protein kinase
CC   inhibitor (CaM-KIIN) and its encoding polynucleotide, variants and
CC   fragments. Also included are an expression system comprising a
CC   polynucleotide capable of producing CaM-KIIN, when the expression vector
CC   is present in a compatible host cell, producing CaM-KIIN, comprising
CC   culturing the cell under expression conditions, and recovering the
CC   polypeptide, a fusion protein consisting of the Immunoglobulin Fc-region
CC   and CaM-KIIN, an anti CaM-KIIN antibody and screening for identifying
CC   compounds that stimulate or inhibit the function or level of CaM-KIIN.
CC   CaM-KIIN and the polynucleotide are useful for treating stroke, head
CC   trauma, multiple sclerosis, Parkinson's disease, Alzheimer's disease,
CC   spinal cord injury, mental disorders, memory deficits, epilepsy,
CC   schizophrenia, bipolar disorders, cancers, ischaemic conditions of the
CC   heart, lung disorders caused by malfunctioning of secretion, diseases
CC   caused by disturbed pancreas secretion behaviour, pain, and disorders of
CC   immune system. CaM-KIIN and the polynucleotide are also useful as
CC   vaccines for inducing immunological response in a mammal. CaM-KIIN is
CC   useful in disease diagnosis and in assays for screening agonistic or
CC   antagonistic compounds. CaM-KIIN is also useful for identifying membrane
CC   bound or soluble receptors. The polynucleotide is useful as diagnostic
CC   reagent for detecting mutations in the associated gene, in the
CC   recombinant production of CaM-KIIN, for chromosome localization studies,
CC   as a valuable tool in tissue expression studies, and in gene therapy. The
CC   antibody is useful for isolating and identifying clones expressing CaM-
CC   KIIN, and for purifying CaM-KIIN. The gene for CaM-KIIN is located on
CC   human chromosome 3 in a region associated with hereditary spastic
CC   paraplegia. The present sequence is the human CaM-KIIN protein
XX
SQ   Sequence 79 AA;

  Query Match             56.0%;  Score 253.5;  DB 5;  Length 79;
  Best Local Similarity   64.1%;  Pred. No. 3.3e-25;
  Matches   50;  Conservative   10;  Mismatches   13;  Indels    5;  Gaps    2;

Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:||| ::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYSEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 5
ABR40377
ID   ABR40377 standard; peptide; 27 AA.
XX
AC   ABR40377;
XX
```

DT 21-JUL-2003 (first entry)
XX
DE Calmodulin kinase II inhibitor SEQ ID NO:6.
XX
KW AC3-I; cardiomyopathy; myocardial contractile dysfunction;
KW myocardial infarction; Calmodulin Kinase II; CaMKII; cardiant.
XX
OS Synthetic.
XX
PN WO2003029428-A2.
XX
PD 10-APR-2003.
XX
PF 01-OCT-2002; 2002WO-US031496.
XX
PR 01-OCT-2001; 2001US-0326576P.
PR 08-OCT-2001; 2001US-0328010P.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Anderson M;
XX
DR WPI; 2003-371993/35.
XX
PT Treating or preventing dilated cardiomyopathy or myocardial contractile
PT dysfunction after myocardial infarction in a subject by administering an
PT inhibitor of Calmodulin Kinase II.
XX
PS Claim 6; Page 57; 58pp; English.
XX
CC The invention relates to a novel method for treating or preventing
CC dilated cardiomyopathy or myocardial contractile dysfunction after
CC myocardial infarction in a subject. The method comprises administering an
CC inhibitor of Calmodulin Kinase II (CaMKII). The method of the invention
CC has cardiant activity. The method is useful for treating or preventing
CC dilated cardiomyopathy or myocardial contractile dysfunction after
CC myocardial infarction in a subject. The present sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 27 AA;

Query Match 29.4%; Score 133; DB 6; Length 27;
Best Local Similarity 96.3%; Pred. No. 6.4e-10;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 42 KRPPKLGQIGRSKRVVIEDDRIDDVLK 68
||||||||||:|||||||||||||||
Db 1 KRPPKLGQIGRAKRVVIEDDRIDDVLK 27

RESULT 6
AAG83014
ID AAG83014 standard; protein; 148 AA.
XX
AC AAG83014;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:3122.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
KW endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
DR N-PSDB; AAH53864.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 18; Page 823; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX
SQ Sequence 148 AA;

Query Match 15.8%; Score 71.5; DB 4; Length 148;
Best Local Similarity 31.1%; Pred. No. 0.76;
Matches 19; Conservative 13; Mismatches 20; Indels 9; Gaps 2;

Qy 21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
:|:: |:| || | :|| | : | |: |: |||: |: : : |: |
Db 57 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 107

Qy 81 P 81
|
Db 108 P 108

RESULT 7
ABP40348
ID ABP40348 standard; protein; 217 AA.
XX
AC ABP40348;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5193.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.

```
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN92893.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT   polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS   Disclosure; SEQ ID NO 5193; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences can
CC   also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life cycle
CC   or inhibit S. epidermidis infection. N.B. The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site
XX
SQ   Sequence 217 AA;

  Query Match             15.8%;  Score 71.5;  DB 5;  Length 217;
  Best Local Similarity   31.1%;  Pred. No. 1.3;
  Matches   19;  Conservative   13;  Mismatches   20;  Indels    9;  Gaps    2;

Qy         21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
              :|:.    |:| ||   | :|| | :      | |:  |: |||:  |: : : |: |
Db        126 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 176

Qy         81 P 81
              |
Db        177 P 177


RESULT 8
ADS04904
ID   ADS04904 standard; protein; 217 AA.
XX
AC   ADS04904;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Staphylococcus epidermis polypeptide seqid 4199.
XX
KW   antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW   recombinant expression vector; infection; computer readable medium;
KW   computer based system.
XX
OS   Staphylococcus epidermidis.
XX
PN   US2004147734-A1.
XX
PD   29-JUL-2004.
XX
PF   01-DEC-2003; 2003US-00724972.
XX
PR   08-NOV-1997;   97US-0064964P.
PR   13-AUG-1998;   98US-00134001.
PR   29-NOV-1999;   99US-00450969.
XX
PA   (DOUC/) DOUCETTE-STAMM L.
PA   (BUSH/) BUSH D.
XX
PI   Doucette-Stamm L,  Bush D;
XX
DR   WPI; 2004-580138/56.
DR   N-PSDB; ADS01132.
XX
PT   New isolated polypeptide and encoding nucleic acid derived from
PT   Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT   treating an S. epidermidis bacterial infection.
XX
PS   Claim 17; SEQ ID NO 4199; 741pp; English.
XX
CC   The invention describes an isolated nucleic acid comprising a nucleotide
CC   sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC   1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC   of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC   given in the specification. Also described are: a recombinant expression
CC   vector; a cell comprising a recombinant expression vector of (1);
CC   producing an S. epidermidis polypeptide; an isolated nucleic acid
CC   comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC   vaccine composition for prevention or treatment of an S. epidermidis
CC   infection, comprising a nucleic acid cited above and a carrier; treating
CC   a subject for S. epidermidis infection; a recombinant or substantially
CC   pure preparation of an S. epidermidis polypeptide or its fragment; a
CC   vaccine composition for prevention or treatment of an S. epidermidis
CC   infection; detecting the presence of a Staphylococcus nucleic acid in a
CC   sample; a computer readable medium having recorded in it the nucleotide
CC   sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC   system for identifying fragments of the Staphylococcus genome of
CC   commercial importance; a computer based system for identifying fragments
CC   of the Staphylococcus plasmids of commercial importance; identifying
CC   commercially important nucleic acid fragments of the Staphylococcus
CC   genome and/or plasmids; and identifying an expression modulating fragment
CC   of the Staphylococcus genome and/or plasmids. The methods and
CC   compositions of the present invention are useful for the diagnosis,
CC   prevention and/or treatment of an Staphylococcal epidermidis bacterial
CC   infection. This is the amino acid sequence of a S. epidermis protein of
CC   the invention.
XX
SQ   Sequence 217 AA;

  Query Match             15.8%;  Score 71.5;  DB 8;  Length 217;
  Best Local Similarity   31.1%;  Pred. No. 1.3;
  Matches   19;  Conservative   13;  Mismatches   20;  Indels    9;  Gaps    2;

Qy         21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
              :|::    |:| ||   | :|| | :      | |:  |: |||:  |: : : |: |
Db        126 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 176

Qy         81 P 81
              |
Db        177 P 177


RESULT 9
AAR10720
ID   AAR10720 standard; protein; 145 AA.
XX
AC   AAR10720;
XX
DT   09-SEP-2004  (revised)
DT   25-MAR-2003  (revised)
DT   09-JAN-2003  (revised)
DT   22-APR-1991  (first entry)
XX
DE   (Gly59)IL-3.
XX
KW   Interleukin-3; mutant; antagonist; cytopenia; immunodeficiency;
KW   immunosuppression; charge reversal.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .12
FT                   /label= signal sequence
FT   Misc-difference 71
FT                   /note= "Glu -> Gly"
XX
PN   EP413383-A.
XX
```

```
PD   20-FEB-1991.
XX
PF   02-AUG-1990;   90EP-00202117.
XX
PR   14-AUG-1989;   89EP-00202082.
PR   14-SEP-1989;   89EP-00202331.
XX
PA   (KONN ) GIST-BROCADES NV.
PA   (DORS/) DORSSERS L C.
XX
PI   Dorssers LCJ,  Vanleen RW;
XX
DR   WPI; 1991-052685/08.
XX
PT   New interleukin-3 mutants - having deletions covering complete coding
PT   sequence while retaining biological activity, have antagonistic effect.
XX
PS   Claim 6; Page 18; 26pp; English.
XX
CC   The mutant was prepd. by recombinant DNA techniques using mutagenic
CC   primers to change the codons of cDNA encoding IL-3. The peptide pref.
CC   also has a deletion at the N-terminus (AAs 1-14) and/or the C-terminus
CC   (AAs 116-133,120-130 and 130-133). The peptide has antagonistic activity
CC   and can be used for therapeutic and diagnostic purposes. See also
CC   AAR10711-R10731 and AAR11051-R11055. (Updated on 09-JAN-2003 to add
CC   missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
CC
CC   Revised record issued on 09-SEP-2004 : Correction to feature table key
XX
SQ   Sequence 145 AA;

  Query Match             15.7%;  Score 71;  DB 2;  Length 145;
  Best Local Similarity   32.2%;  Pred. No. 0.87;
  Matches   19;  Conservative    8;  Mismatches   26;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  | |      | | | ||   |: : :     |: :|||:    || : :|
Db         46 LLDFNNLNGEDQDILMENNLRRPNLGAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 104


RESULT 10
ABU50950
ID   ABU50950 standard; protein; 140 AA.
XX
AC   ABU50950;
XX
DT   07-MAY-2003  (first entry)
XX
DE   Helicobacter pylori selected interacting domain (SID) protein #293.
XX
KW   Protein-protein interaction; ulcer; selected interacting domain; SID.
XX
OS   Helicobacter pylori.
XX
PN   WO200266501-A2.
XX
PD   29-AUG-2002.
XX
PF   28-DEC-2001; 2001WO-EP015428.
XX
PR   02-JAN-2001; 2001US-0259302P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (INSP ) INST PASTEUR.
XX
PI   Legrain P,  Rain J,  Colland F,  De Reuse H,  Labigne A;
XX
DR   WPI; 2002-674910/72.
DR   N-PSDB; ABX65694.
XX
PT   New complexes of protein-protein interactions in Helicobacter pylori,
PT   useful for identifying modulating compounds for treating or preventing
PT   ulcers in mammals.
XX
PS   Claim 6; Page 157; 642pp; English.
XX
CC   The invention describes a complex of protein-protein interactions in
CC   Helicobacter pylori selected from 421 complexes given in the
CC   specification. The complex of protein-protein interactions are useful for
CC   screening for agents which modulate the interaction of proteins.
CC   Modulating compounds which binds to a targeted bacterial protein may be
CC   used for treating or preventing ulcers in a human or animal. This is the
CC   amino acid sequence of a selected interacting domain (SID), identified
CC   via protein-protein interactions. Note: Where the patent number printed
CC   at the top of the pages in the specification has obscured areas of
CC   protein sequence, the indexer has replaced the residue with an X to
CC   represent an illegible residue
XX
SQ   Sequence 140 AA;

  Query Match             15.1%;  Score 68.5;  DB 5;  Length 140;
  Best Local Similarity   23.3%;  Pred. No. 1.8;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                 |||   | |:   |:|    :| :   ::
Db         59 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 102

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db        103 DILNQMIEQAPLINKLISWLGPQDLS 128


RESULT 11
AAB52498
ID   AAB52498 standard; protein; 142 AA.
XX
AC   AAB52498;
XX
DT   23-FEB-2001  (first entry)
XX
DE   Helicobacter pylori bait polypeptide #16.
XX
KW   Helicobacter pylori; two-hybrid system; protein-protein interaction;
KW   bait polypeptide; gastric ulcer; antibacterial.
XX
OS   Helicobacter pylori.
XX
PN   WO200066722-A1.
XX
PD   09-NOV-2000.
XX
PF   14-APR-2000; 2000WO-IB000603.
XX
PR   30-APR-1999;   99EP-00401066.
XX
PA   (HYBR-) HYBRIGENICS SA.
XX
PI   Legrain P,  Selig L,  Rain J;
XX
DR   WPI; 2000-687535/67.
DR   N-PSDB; AAC97244.
XX
PT   A two-hybrid system for identifying compounds useful in the treatment of
PT   e.g. gastric ulcers comprises producing a collection of recombinant cell
PT   clones.
XX
PS   Example 5; Page 98; 267pp; English.
XX
CC   The present sequence is a bait polypeptide used in a Helicobacter pylori
CC   two-hybrid screen to identify protein-protein interactions. The method is
CC   used to identify a recombinant cell clone expressing a prey polypeptide
CC   which is capable of interacting with the bait polypeptide. The two hybrid
CC   system is useful for screening compounds for antibacterial activity. It
CC   may be used in the treatment of gastric ulcers. The polynucleotides are
CC   useful as amplification primers or specific detection probes. The
```

```
CC   polypeptides, vectors or host cells can be used as immunogens to produce
CC   mono- or polyclonal antibodies. The polynucleotides, polypeptides,
CC   antibodies, vectors, host cells or modulating agents can be used to
CC   produce a pharmaceutical composition
XX
SQ   Sequence 142 AA;

  Query Match             15.1%;  Score 68.5;  DB 3;  Length 142;
  Best Local Similarity   23.3%;  Pred. No. 1.8;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db         59 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 102

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db        103 DILNQMIEQAPLINKLISWLGPQDLS 128


RESULT 12
ABU51049
ID   ABU51049 standard; protein; 155 AA.
XX
AC   ABU51049;
XX
DT   07-MAY-2003  (first entry)
XX
DE   Helicobacter pylori selected interacting domain (SID) protein #392.
XX
KW   Protein-protein interaction; ulcer; selected interacting domain; SID.
XX
OS   Helicobacter pylori.
XX
PN   WO200266501-A2.
XX
PD   29-AUG-2002.
XX
PF   28-DEC-2001; 2001WO-EP015428.
XX
PR   02-JAN-2001; 2001US-0259302P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (INSP ) INST PASTEUR.
XX
PI   Legrain P,  Rain J,  Colland F,  De Reuse H,  Labigne A;
XX
DR   WPI; 2002-674910/72.
DR   N-PSDB; ABX65793.
XX
PT   New complexes of protein-protein interactions in Helicobacter pylori,
PT   useful for identifying modulating compounds for treating or preventing
PT   ulcers in mammals.
XX
PS   Claim 6; Page 182; 642pp; English.
XX
CC   The invention describes a complex of protein-protein interactions in
CC   Helicobacter pylori selected from 421 complexes given in the
CC   specification. The complex of protein-protein interactions are useful for
CC   screening for agents which modulate the interaction of proteins.
CC   Modulating compounds which binds to a targeted bacterial protein may be
CC   used for treating or preventing ulcers in a human or animal. This is the
CC   amino acid sequence of a selected interacting domain (SID), identified
CC   via protein-protein interactions. Note: Where the patent number printed
CC   at the top of the pages in the specification has obscured areas of
CC   protein sequence, the indexer has replaced the residue with an X to
CC   represent an illegible residue
XX
SQ   Sequence 155 AA;

  Query Match             15.1%;  Score 68.5;  DB 5;  Length 155;
  Best Local Similarity   23.3%;  Pred. No. 2;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db          1 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 44

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db         45 DILNQMIEQAPLINKLISWLGPQDLS 70


RESULT 13
AAB52603
ID   AAB52603 standard; protein; 157 AA.
XX
AC   AAB52603;
XX
DT   23-FEB-2001  (first entry)
XX
DE   Helicobacter pylori bait polypeptide #121.
XX
KW   Helicobacter pylori; two-hybrid system; protein-protein interaction;
KW   bait polypeptide; gastric ulcer; antibacterial.
XX
OS   Helicobacter pylori.
XX
PN   WO200066722-A1.
XX
PD   09-NOV-2000.
XX
PF   14-APR-2000; 2000WO-IB000603.
XX
PR   30-APR-1999;   99EP-00401066.
XX
PA   (HYBR-) HYBRIGENICS SA.
XX
PI   Legrain P,  Selig L,  Rain J;
XX
DR   WPI; 2000-687535/67.
DR   N-PSDB; AAC97349.
XX
PT   A two-hybrid system for identifying compounds useful in the treatment of
PT   e.g. gastric ulcers comprises producing a collection of recombinant cell
PT   clones.
XX
PS   Example 5; Page 233-234; 267pp; English.
XX
CC   The present sequence is a bait polypeptide used in a Helicobacter pylori
CC   two-hybrid screen to identify protein-protein interactions. The method is
CC   used to identify a recombinant cell clone expressing a prey polypeptide
CC   which is capable of interacting with the bait polypeptide. The two hybrid
CC   system is useful for screening compounds for antibacterial activity. It
CC   may be used in the treatment of gastric ulcers. The polynucleotides are
CC   useful as amplification primers or specific detection probes. The
CC   polypeptides, vectors or host cells can be used as immunogens to produce
CC   mono- or polyclonal antibodies. The polynucleotides, polypeptides,
CC   antibodies, vectors, host cells or modulating agents can be used to
CC   produce a pharmaceutical composition
XX
SQ   Sequence 157 AA;

  Query Match             15.1%;  Score 68.5;  DB 3;  Length 157;
  Best Local Similarity   23.3%;  Pred. No. 2.1;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db          1 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 44

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db         45 DILNQMIEQAPLINKLISWLGPQDLS 70
```

```
RESULT 14
ABU52122
ID   ABU52122 standard; protein; 289 AA.
XX
AC   ABU52122;
XX
DT   07-MAY-2003  (first entry)
XX
DE   Helicobacter pylori selected interacting domain (SID) protein #1466.
XX
KW   Protein-protein interaction; ulcer; selected interacting domain; SID.
XX
OS   Helicobacter pylori.
XX
PN   WO200266501-A2.
XX
PD   29-AUG-2002.
XX
PF   28-DEC-2001; 2001WO-EP015428.
XX
PR   02-JAN-2001; 2001US-0259302P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (INSP ) INST PASTEUR.
XX
PI   Legrain P,  Rain J,  Colland F,  De Reuse H,  Labigne A;
XX
DR   WPI; 2002-674910/72.
DR   N-PSDB; ABX66867.
XX
PT   New complexes of protein-protein interactions in Helicobacter pylori,
PT   useful for identifying modulating compounds for treating or preventing
PT   ulcers in mammals.
XX
PS   Claim 6; Page 441; 642pp; English.
XX
CC   The invention describes a complex of protein-protein interactions in
CC   Helicobacter pylori selected from 421 complexes given in the
CC   specification. The complex of protein-protein interactions are useful for
CC   screening for agents which modulate the interaction of proteins.
CC   Modulating compounds which binds to a targeted bacterial protein may be
CC   used for treating or preventing ulcers in a human or animal. This is the
CC   amino acid sequence of a selected interacting domain (SID), identified
CC   via protein-protein interactions. Note: Where the patent number printed
CC   at the top of the pages in the specification has obscured areas of
CC   protein sequence, the indexer has replaced the residue with an X to
CC   represent an illegible residue
XX
SQ   Sequence 289 AA;

  Query Match             15.1%;  Score 68.5;  DB 5;  Length 289;
  Best Local Similarity   23.3%;  Pred. No. 4.5;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db         75 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 118

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db        119 DILNQMIEQAPLINKLISWLGPQDLS 144


RESULT 15
ABU51075
ID   ABU51075 standard; protein; 659 AA.
XX
AC   ABU51075;
XX
DT   07-MAY-2003  (first entry)
XX
DE   Helicobacter pylori selected interacting domain (SID) protein #418.
XX
KW   Protein-protein interaction; ulcer; selected interacting domain; SID.
XX
OS   Helicobacter pylori.
XX
PN   WO200266501-A2.
XX
PD   29-AUG-2002.
XX
PF   28-DEC-2001; 2001WO-EP015428.
XX
PR   02-JAN-2001; 2001US-0259302P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (INSP ) INST PASTEUR.
XX
PI   Legrain P,  Rain J,  Colland F,  De Reuse H,  Labigne A;
XX
DR   WPI; 2002-674910/72.
DR   N-PSDB; ABX65819.
XX
PT   New complexes of protein-protein interactions in Helicobacter pylori,
PT   useful for identifying modulating compounds for treating or preventing
PT   ulcers in mammals.
XX
PS   Claim 6; Page 188; 642pp; English.
XX
CC   The invention describes a complex of protein-protein interactions in
CC   Helicobacter pylori selected from 421 complexes given in the
CC   specification. The complex of protein-protein interactions are useful for
CC   screening for agents which modulate the interaction of proteins.
CC   Modulating compounds which binds to a targeted bacterial protein may be
CC   used for treating or preventing ulcers in a human or animal. This is the
CC   amino acid sequence of a selected interacting domain (SID), identified
CC   via protein-protein interactions. Note: Where the patent number printed
CC   at the top of the pages in the specification has obscured areas of
CC   protein sequence, the indexer has replaced the residue with an X to
CC   represent an illegible residue
XX
SQ   Sequence 659 AA;

  Query Match             15.1%;  Score 68.5;  DB 5;  Length 659;
  Best Local Similarity   23.3%;  Pred. No. 13;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db        422 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 465

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db        466 DILNQMIEQAPLINKLISWLGPQDLS 491


Search completed: March 28, 2005, 12:59:07
Job time : 128 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 28, 2005, 11:19:30 ; Search time 53 Seconds
(without alignments)
119.720 Million cell updates/sec

Title: US-10-071-645-2
Perfect score: 453
Sequence: 1 MWEVLPYGDEKLSPYGDGGD.....VLKNMTDKAPLVSNSPKTMS 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71.5	15.8	148	4	US-09-710-279-3122	Sequence 3122, Ap
2	71.5	15.8	217	3	US-09-134-001C-5193	Sequence 5193, Ap
3	68	15.0	112	1	US-08-081-539-72	Sequence 72, Appl
4	68	15.0	112	1	US-08-466-647-72	Sequence 72, Appl
5	67	14.8	114	1	US-08-081-539-64	Sequence 64, Appl
6	67	14.8	114	1	US-08-466-647-64	Sequence 64, Appl
7	66	14.6	105	1	US-08-081-539-73	Sequence 73, Appl
8	66	14.6	105	1	US-08-466-647-73	Sequence 73, Appl
9	65	14.3	301	1	US-07-797-553-2	Sequence 2, Appli
10	65	14.3	303	1	US-07-797-553-4	Sequence 4, Appli
11	64.5	14.2	248	4	US-09-252-991A-19495	Sequence 19495, A
12	63.5	14.0	251	4	US-09-252-991A-26403	Sequence 26403, A
13	63	13.9	110	1	US-08-081-539-69	Sequence 69, Appl
14	63	13.9	110	1	US-08-466-647-69	Sequence 69, Appl
15	63	13.9	111	1	US-08-411-795B-65	Sequence 65, Appl
16	63	13.9	111	1	US-08-411-795B-74	Sequence 74, Appl
17	63	13.9	111	1	US-08-411-795B-75	Sequence 75, Appl
18	63	13.9	111	1	US-08-411-795B-407	Sequence 407, App
19	63	13.9	111	1	US-08-411-796-23	Sequence 23, Appl
20	63	13.9	111	1	US-08-411-796-65	Sequence 65, Appl
21	63	13.9	111	1	US-08-411-796-217	Sequence 217, App
22	63	13.9	111	1	US-08-411-796-218	Sequence 218, App
23	63	13.9	111	2	US-08-469-319A-65	Sequence 65, Appl
24	63	13.9	111	2	US-08-469-319A-74	Sequence 74, Appl
25	63	13.9	111	2	US-08-469-319A-75	Sequence 75, Appl
26	63	13.9	111	2	US-08-469-319A-407	Sequence 407, App
27	63	13.9	111	3	US-08-471-039-23	Sequence 23, Appl
28	63	13.9	111	3	US-08-471-039-65	Sequence 65, Appl
29	63	13.9	111	3	US-08-471-039-217	Sequence 217, App
30	63	13.9	111	3	US-08-471-039-218	Sequence 218, App
31	63	13.9	111	3	US-08-469-318-17	Sequence 17, Appl
32	63	13.9	111	3	US-08-469-318-18	Sequence 18, Appl
33	63	13.9	111	3	US-08-468-609A-17	Sequence 17, Appl
34	63	13.9	111	3	US-08-468-609A-18	Sequence 18, Appl
35	63	13.9	111	3	US-08-192-325B-17	Sequence 17, Appl
36	63	13.9	111	3	US-08-192-325B-18	Sequence 18, Appl
37	63	13.9	111	3	US-08-471-045-17	Sequence 17, Appl
38	63	13.9	111	3	US-08-471-045-18	Sequence 18, Appl
39	63	13.9	111	3	US-08-469-712A-17	Sequence 17, Appl
40	63	13.9	111	3	US-08-469-712A-18	Sequence 18, Appl
41	63	13.9	111	3	US-08-193-373A-17	Sequence 17, Appl
42	63	13.9	111	3	US-08-193-373A-18	Sequence 18, Appl
43	63	13.9	111	3	US-08-446-871-17	Sequence 17, Appl
44	63	13.9	111	3	US-08-446-871-18	Sequence 18, Appl
45	63	13.9	111	3	US-08-446-872A-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-710-279-3122
; Sequence 3122, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3122
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-3122

Query Match 15.8%; Score 71.5; DB 4; Length 148;
Best Local Similarity 31.1%; Pred. No. 0.12;
Matches 19; Conservative 13; Mismatches 20; Indels 9; Gaps 2;

```
Qy         21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
              :|::    |:| ||   | :|| | :      | |:  |: |||:  |: : : |: |
Db         57 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 107

Qy         81 P 81
              |
Db        108 P 108
```

RESULT 2
US-09-134-001C-5193
; Sequence 5193, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779

```
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5193
;   LENGTH: 217
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5193

  Query Match             15.8%;  Score 71.5;  DB 3;  Length 217;
  Best Local Similarity   31.1%;  Pred. No. 0.2;
  Matches   19;  Conservative   13;  Mismatches   20;  Indels    9;  Gaps    2;

Qy         21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
              :|::    |:| ||   | :||  | :      | |:  |: |||:  |: : :|: |
Db        126 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 176

Qy         81 P 81
              |
Db        177 P 177


RESULT 3
US-08-081-539-72
; Sequence 72, Application US/08081539
; Patent No. 5501962
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,539
;      FILING DATE:  19930621
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  112 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-081-539-72

  Query Match             15.0%;  Score 68;  DB 1;  Length 112;
  Best Local Similarity   33.9%;  Pred. No. 0.25;
  Matches   20;  Conservative    6;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  | |      | | | |    ||:  |     |: :|||:    || : :|
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRSQGEVDNASAIESILKNLLPCLPLATAAP 79


RESULT 4
US-08-466-647-72
; Sequence 72, Application US/08466647
; Patent No. 5543141
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,647
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,539
;      FILING DATE:  21-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  112 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-466-647-72

  Query Match             15.0%;  Score 68;  DB 1;  Length 112;
  Best Local Similarity   33.9%;  Pred. No. 0.25;
  Matches   20;  Conservative    6;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  | |      | | | |    ||:  |     |: :|||:    || : :|
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRSQGEVDNASAIESILKNLLPCLPLATAAP 79


RESULT 5
US-08-081-539-64
; Sequence 64, Application US/08081539
; Patent No. 5501962
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,539
;      FILING DATE:  19930621
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  64:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  114 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-081-539-64

  Query Match             14.8%;  Score 67;  DB 1;  Length 114;
  Best Local Similarity   32.8%;  Pred. No. 0.35;
  Matches   20;  Conservative    9;  Mismatches   24;  Indels    8;  Gaps    2;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDR--IDDVLKNMTDKAPLVSNS 80
              | | ||  |  |      | | | |    |: : :  :||  |: :|||:    || : :
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQPEDRYVIESILKNLLPCLPLATAA 80

Qy         81 P 81
              |
Db         81 P 81


RESULT 6
US-08-466-647-64
; Sequence 64, Application US/08466647
; Patent No. 5543141
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,647
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,539
;      FILING DATE:  21-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  64:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  114 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-466-647-64

  Query Match             14.8%;  Score 67;  DB 1;  Length 114;
  Best Local Similarity   32.8%;  Pred. No. 0.35;
  Matches   20;  Conservative    9;  Mismatches   24;  Indels    8;  Gaps    2;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDR--IDDVLKNMTDKAPLVSNS 80
              | | ||  |  |      | | | |    |: : :  :||  |: :|||:    || : :
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQPEDRYVIESILKNLLPCLPLATAA 80

Qy         81 P 81
              |
Db         81 P 81


RESULT 7
US-08-081-539-73
; Sequence 73, Application US/08081539
; Patent No. 5501962
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,539
;      FILING DATE:  19930621
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  73:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-081-539-73

  Query Match             14.6%;  Score 66;  DB 1;  Length 105;
  Best Local Similarity   30.5%;  Pred. No. 0.43;
  Matches   18;  Conservative    9;  Mismatches   26;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              |:| ||  |  |       | | | |    |: : :     |: :|||:    || : :|
Db         14 LKDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 72


RESULT 8
US-08-466-647-73
; Sequence 73, Application US/08466647
; Patent No. 5543141
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,647
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,539
;      FILING DATE:  21-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  73:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-466-647-73

  Query Match             14.6%;  Score 66;  DB 1;  Length 105;
  Best Local Similarity   30.5%;  Pred. No. 0.43;
  Matches   18;  Conservative    9;  Mismatches   26;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              |:| ||  |  |       | | | |    |: : :     |: :|||:    || : :|
Db         14 LKDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 72


RESULT 9
US-07-797-553-2
; Sequence 2, Application US/07797553
; Patent No. 5376367
;  GENERAL INFORMATION:
;    APPLICANT:  Williams, Douglas E.
;    TITLE OF INVENTION:  Fusion Proteins Comprising MGF and IL-3
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/797,553
;      FILING DATE:  19911122
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wight, Christopher L.
;      REGISTRATION NUMBER:  31,680
;      REFERENCE/DOCKET NUMBER:  0525
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-587-0430
;      TELEFAX:  206-587-0606
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  301 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-797-553-2

  Query Match             14.3%;  Score 65;  DB 1;  Length 301;
  Best Local Similarity   30.5%;  Pred. No. 2.3;
  Matches   18;  Conservative    9;  Mismatches   26;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  |  |       | | | |    |: : : :   |: :|||:    || : :|
Db         34 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQDASAIESILKNLLPCLPLATAAP 92


RESULT 10
US-07-797-553-4
; Sequence 4, Application US/07797553
; Patent No. 5376367
;  GENERAL INFORMATION:
;    APPLICANT:  Williams, Douglas E.
;    TITLE OF INVENTION:  Fusion Proteins Comprising MGF and IL-3
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/797,553
;      FILING DATE:  19911122
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wight, Christopher L.
;      REGISTRATION NUMBER:  31,680
;      REFERENCE/DOCKET NUMBER:  0525
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-587-0430
;      TELEFAX:  206-587-0606
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-797-553-4

  Query Match             14.3%;  Score 65;  DB 1;  Length 303;
  Best Local Similarity   30.5%;  Pred. No. 2.3;
  Matches   18;  Conservative    9;  Mismatches   26;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  | |       | | | |    |: : : :   |: :|||:    || : :|
Db        204 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQDASAIESILKNLLPCLPLATAAP 262


RESULT 11
US-09-252-991A-19495
; Sequence 19495, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 19495
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19495

  Query Match             14.2%;  Score 64.5;  DB 4;  Length 248;
  Best Local Similarity   39.0%;  Pred. No. 2.1;
  Matches   23;  Conservative    7;  Mismatches   22;  Indels    7;  Gaps    4;

Qy         13 SPYGDGGD---VGQ--IFSCRLQDTNNFFGAG-QNKRP-PKLGQIGRSKRVVIEDDRID 64
              |  | |||   |||  :   | | |:     | | :|| | | |:|:    : || |:|
Db         96 SQEGSGGDHHQVGQQALQGPRRQGTDRAQAKGRQGQRPEPGLRQVGQLAGGIAEDQRVD 154


RESULT 12
US-09-252-991A-26403
; Sequence 26403, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26403
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26403

  Query Match             14.0%;  Score 63.5;  DB 4;  Length 251;
  Best Local Similarity   28.2%;  Pred. No. 2.9;
  Matches   24;  Conservative   13;  Mismatches   31;  Indels   17;  Gaps    3;

Qy         12 LSPYGDG-GDVGQIFSCRLQDTNNFFGAGQN---------KRPPKLGQIGRSK------- 54
              |  :|||     |    ||     |   |::         :|   : :|||:
Db        152 LGDPGDGQSGAEQTDDFRLPGRQPFVVRGKHLVDEDVPDRRREILVSRIGRAHRRQQVTG 211

Qy         55 RVVIEDDRIDDVLKNMTDKAPLVSN 79
              || : :| || :|:|  ||| :| :
Db        212 RVGLAEDAIDTLLENGPDKARIVQH 236


RESULT 13
US-08-081-539-69
; Sequence 69, Application US/08081539
; Patent No. 5501962
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,539
;      FILING DATE:  19930621
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  110 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-081-539-69

  Query Match             13.9%;  Score 63;  DB 1;  Length 110;
  Best Local Similarity   30.5%;  Pred. No. 1.2;
  Matches   18;  Conservative    8;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  | |       | | | |    |: : :     |: :|||:    || : :|
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 79


RESULT 14
US-08-466-647-69
; Sequence 69, Application US/08466647
; Patent No. 5543141
;  GENERAL INFORMATION:
;    APPLICANT:  Braford-Goldberg, Sarah R.
```

```
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Human/Murine
;    TITLE OF INVENTION:  Chimeric Hybrid Polypeptides
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mary Jo Kanady, G. D. Searle & Co., Corporate
;      ADDRESSEE:  Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,647
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,539
;      FILING DATE:  21-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kanady, Mary J.
;      REGISTRATION NUMBER:  28623
;      REFERENCE/DOCKET NUMBER:  2724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  110 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-466-647-69

  Query Match             13.9%;  Score 63;  DB 1;  Length 110;
  Best Local Similarity   30.5%;  Pred. No. 1.2;
  Matches   18;  Conservative    8;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  |  |      | | | |    |: : :     |: :|||:    || : :|
Db         21 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 79


RESULT 15
US-08-411-795B-65
; Sequence 65, Application US/08411795B
; Patent No. 5604116
;  GENERAL INFORMATION:
;    APPLICANT:  Abrams, Mark A.
;    APPLICANT:  Bauer, S. C.
;    APPLICANT:  Braford-Goldberg, Sarah R.
;    APPLICANT:  Caparon, Maire H.
;    APPLICANT:  Easton, Alan M.
;    APPLICANT:  Klein, Barbara K.
;    APPLICANT:  McKearn, John P.
;    APPLICANT:  Olins, Peter O.
;    APPLICANT:  Paik, Kumnan
;    APPLICANT:  Thomas, John W.
;    TITLE OF INVENTION:  Interleukin-3 (IL-3) Multiple Mutation
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  415
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis A. Bennett, G.D. Searle & Co.,
;      ADDRESSEE:  Corporate Patent Dept.
;      STREET:  P. O. Box 5110
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/411,795B
;      FILING DATE:  04-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/981,044
;      FILING DATE:  24-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/11197
;      FILING DATE:  22-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bennett, Dennis A.
;      REGISTRATION NUMBER:  34,547
;      REFERENCE/DOCKET NUMBER:  C2713/2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708)470-6501
;      TELEFAX:  (708)470-6881
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-411-795B-65

  Query Match             13.9%;  Score 63;  DB 1;  Length 111;
  Best Local Similarity   30.5%;  Pred. No. 1.2;
  Matches   18;  Conservative    8;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  |  |      | | | |    |: : :     |: :|||:    || : :|
Db         20 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 78


Search completed: March 28, 2005, 13:03:33
Job time : 55 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 28, 2005, 13:01:39 ; Search time 1199 Seconds
(without alignments)
23.473 Million cell updates/sec

Title: US-10-071-645-2
Perfect score: 453
Sequence: 1 MWEVLPYGDEKLSPYGDGGD.....VLKNMTDKAPLVSNSPKTMS 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	453	100.0	85	14	US-10-071-645-2	Sequence 2, Appli
2	391	86.3	78	15	US-10-136-728-22	Sequence 22, Appl
3	253.5	56.0	79	15	US-10-362-674-2	Sequence 2, Appli
4	73	16.1	154	16	US-10-767-701-39490	Sequence 39490, A
5	72.5	16.0	770	16	US-10-437-963-137267	Sequence 137267,
6	70	15.5	116	15	US-10-424-599-222962	Sequence 222962,
7	68.5	15.1	142	14	US-10-012-819-32	Sequence 32, Appl
8	68.5	15.1	157	14	US-10-012-819-242	Sequence 242, App
9	68.5	15.1	1943	15	US-10-282-122A-58750	Sequence 58750, A
10	68.5	15.1	1974	9	US-09-895-913A-12	Sequence 12, Appl
11	67.5	14.9	1120	15	US-10-335-977-5265	Sequence 5265, Ap
12	67.5	14.9	1576	15	US-10-335-977-5266	Sequence 5266, Ap
13	67.5	14.9	1775	15	US-10-335-977-5267	Sequence 5267, Ap
14	67.5	14.9	3194	15	US-10-282-122A-59046	Sequence 59046, A
15	67	14.8	932	15	US-10-369-493-3737	Sequence 3737, Ap
16	66	14.6	575	15	US-10-369-493-12520	Sequence 12520, A
17	65	14.3	133	16	US-10-658-834A-478	Sequence 478, App
18	65	14.3	133	16	US-10-658-834A-485	Sequence 485, App
19	65	14.3	946	16	US-10-437-963-180294	Sequence 180294,
20	65	14.3	958	16	US-10-437-963-180295	Sequence 180295,
21	64.5	14.2	185	16	US-10-767-701-56453	Sequence 56453, A
22	64	14.1	133	16	US-10-658-834A-477	Sequence 477, App
23	63.5	14.0	652	16	US-10-437-963-183833	Sequence 183833,
24	63.5	14.0	781	15	US-10-424-599-265840	Sequence 265840,
25	63.5	14.0	988	16	US-10-437-963-123768	Sequence 123768,
26	63.5	14.0	1445	15	US-10-149-310-110	Sequence 110, App
27	63	13.9	111	14	US-10-090-182A-65	Sequence 65, Appl
28	63	13.9	111	14	US-10-090-182A-74	Sequence 74, Appl
29	63	13.9	111	14	US-10-090-182A-75	Sequence 75, Appl
30	63	13.9	111	14	US-10-090-182A-407	Sequence 407, App
31	63	13.9	111	14	US-10-083-446-17	Sequence 17, Appl
32	63	13.9	111	14	US-10-083-446-18	Sequence 18, Appl
33	63	13.9	111	14	US-10-072-571-17	Sequence 17, Appl
34	63	13.9	111	14	US-10-072-571-18	Sequence 18, Appl
35	63	13.9	111	15	US-10-078-113-65	Sequence 65, Appl
36	63	13.9	111	15	US-10-078-113-74	Sequence 74, Appl
37	63	13.9	111	15	US-10-078-113-75	Sequence 75, Appl
38	63	13.9	111	15	US-10-078-113-407	Sequence 407, App
39	63	13.9	111	15	US-10-179-940-23	Sequence 23, Appl
40	63	13.9	111	15	US-10-179-940-65	Sequence 65, Appl
41	63	13.9	111	15	US-10-179-940-217	Sequence 217, App
42	63	13.9	111	15	US-10-179-940-218	Sequence 218, App
43	63	13.9	111	17	US-10-948-576-17	Sequence 17, Appl
44	63	13.9	111	17	US-10-948-576-18	Sequence 18, Appl
45	63	13.9	112	14	US-10-090-182A-138	Sequence 138, App

ALIGNMENTS

RESULT 1
US-10-071-645-2
; Sequence 2, Application US/10071645
; Publication No. US20030148389A1
; GENERAL INFORMATION:
; APPLICANT: Bihain, Bernard
; APPLICANT: Bour, Barbara
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Schizophrenia Related Gene and Protein
; FILE REFERENCE: 92.US2.CIP
; CURRENT APPLICATION NUMBER: US/10/071,645
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/223,482
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01891
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 2
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-071-645-2

Query Match 100.0%; Score 453; DB 14; Length 85;
Best Local Similarity 100.0%; Pred. No. 5.7e-50;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAPLVSNSPKTMS 85
              |||||||||||||||||||||||||
Db         61 DRIDDVLKNMTDKAPLVSNSPKTMS 85
```

```
RESULT 2
US-10-136-728-22
; Sequence 22, Application US/10136728
; Publication No. US20030236188A1
; GENERAL INFORMATION:
;  APPLICANT: Spytek, Kimberly A.
;  APPLICANT:  Li, Li
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Anderson, David W.
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gerlach, Valerie L.
;  APPLICANT:  Taupier, Raymond J.
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Perman, John A.
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: No. US20030236188A1el Human Proteins, Polynucleotides Encoding Th
;  TITLE OF INVENTION:  The Same
;  FILE REFERENCE: 21402-347 D (Cura 647 Other)
;  CURRENT APPLICATION NUMBER: US/10/136,728
;  CURRENT FILING DATE:  2002-05-01
;  PRIOR APPLICATION NUMBER: 60/288,395
;  PRIOR FILING DATE: 2001-05-03
;  PRIOR APPLICATION NUMBER: 60/289,087
;  PRIOR FILING DATE: 2001-05-07
;  PRIOR APPLICATION NUMBER: 60/289,619
;  PRIOR FILING DATE: 2001-05-08
;  PRIOR APPLICATION NUMBER: 60/289,818
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 60/289,817
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 60/290,194
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: 60/290,753
;  PRIOR FILING DATE: 2001-05-14
;  PRIOR APPLICATION NUMBER: 60/291,189
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: 60/292,374
;  PRIOR FILING DATE: 2001-05-21
;  PRIOR APPLICATION NUMBER: 60/293,107
;  PRIOR FILING DATE: 2001-05-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 22
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-136-728-22

  Query Match             86.3%;  Score 391;  DB 15;  Length 78;
  Best Local Similarity   98.7%;  Pred. No. 4.4e-42;
  Matches   74;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAP 75
              ||||||||||||||
Db         61 DRIDDVLKNMTDKAP 75


RESULT 3
US-10-362-674-2
; Sequence 2, Application US/10362674
; Publication No. US20040086973A1
; GENERAL INFORMATION:
;  APPLICANT: Merck Patent GmbH
;  TITLE OF INVENTION: Calmodulin dependent kinase inhibitor
;  FILE REFERENCE: CaMKIINKDWS
;  CURRENT APPLICATION NUMBER: US/10/362,674
;  CURRENT FILING DATE:  2003-08-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-362-674-2

  Query Match             56.0%;  Score 253.5;  DB 15;  Length 79;
  Best Local Similarity   64.1%;  Pred. No. 1.8e-24;
  Matches   50;  Conservative   10;  Mismatches   13;  Indels    5;  Gaps    2;

Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:||| ::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYSEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 4
US-10-767-701-39490
; Sequence 39490, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39490
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-C15416_1.pep
US-10-767-701-39490

  Query Match             16.1%;  Score 73;  DB 16;  Length 154;
  Best Local Similarity   23.6%;  Pred. No. 0.53;
  Matches   21;  Conservative   15;  Mismatches   37;  Indels   16;  Gaps    3;

Qy          2 WEVL---PYGDEKLSPYGDGG------DVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGR 52
              ||:|   |       ||  ||      | | |         :::|  | : : |:   : :
Db          5 WELLLASPKAAVAAEPYAGGGGGGGEDDAGAI-------KHDYFDLGSDAKYPRRASLSK 57

Qy         53 SKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
                | |:: :|: |   :| :  |   |
Db         58 GNDEVDEEEGVDEGLLGASDNSSWVEPDP 86


RESULT 5
US-10-437-963-137267
; Sequence 137267, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
```

```
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137267
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_38766C.1.pep
US-10-437-963-137267

  Query Match             16.0%;  Score 72.5;  DB 16;  Length 770;
  Best Local Similarity   28.8%;  Pred. No. 4.8;
  Matches   17;  Conservative   11;  Mismatches   22;  Indels    9;  Gaps    1;

Qy         23 QIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              || :|   | | |||         :  :   | |::  |  ::::||: |:  |  | |
Db        716 QINNCMHIDNNGFFG---------VSSLPSLKEVMLMGDHNEELMKNLRDQLALNQNQP 765


RESULT 6
US-10-424-599-222962
; Sequence 222962, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 222962
;   LENGTH: 116
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_43364C.1.pep
US-10-424-599-222962

  Query Match             15.5%;  Score 70;  DB 15;  Length 116;
  Best Local Similarity   39.2%;  Pred. No. 0.89;
  Matches   20;  Conservative    7;  Mismatches   22;  Indels    2;  Gaps    2;

Qy          6 PYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPK-LGQIGRSKR 55
              |:| :|  |:| ||  | : | | :  ||  | |   || |   | | :|:
Db          6 PFGKKKGVPWGGGGKQG-LSSPRKKGYNNQTGRGDQPRPKKFFTQPGPTKK 55


RESULT 7
US-10-012-819-32
; Sequence 32, Application US/10012819
; Publication No. US20030017478A1
; GENERAL INFORMATION:
;  APPLICANT: Legrain, Pierre
;  APPLICANT:  Selig, Luc
;  APPLICANT:  Rain, Jean-Christophe
;  TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
;  TITLE OF INVENTION:  pylori Protein-Protein Interactions and Applications thereof
;  FILE REFERENCE: B5053
;  CURRENT APPLICATION NUMBER: US/10/012,819
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: EP 99401066.8
;  PRIOR FILING DATE: 1999-04-30
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
;   LENGTH: 142
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-10-012-819-32

  Query Match             15.1%;  Score 68.5;  DB 14;  Length 142;
  Best Local Similarity   23.3%;  Pred. No. 1.8;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db         59 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 102

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db        103 DILNQMIEQAPLINKLISWLGPQDLS 128


RESULT 8
US-10-012-819-242
; Sequence 242, Application US/10012819
; Publication No. US20030017478A1
; GENERAL INFORMATION:
;  APPLICANT: Legrain, Pierre
;  APPLICANT:  Selig, Luc
;  APPLICANT:  Rain, Jean-Christophe
;  TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
;  TITLE OF INVENTION:  pylori Protein-Protein Interactions and Applications thereof
;  FILE REFERENCE: B5053
;  CURRENT APPLICATION NUMBER: US/10/012,819
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: EP 99401066.8
;  PRIOR FILING DATE: 1999-04-30
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 242
;   LENGTH: 157
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-10-012-819-242

  Query Match             15.1%;  Score 68.5;  DB 14;  Length 157;
  Best Local Similarity   23.3%;  Pred. No. 2;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db          1 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 44

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db         45 DILNQMIEQAPLINKLISWLGPQDLS 70


RESULT 9
US-10-282-122A-58750
; Sequence 58750, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
```

```
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58750
;   LENGTH: 1943
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-10-282-122A-58750

  Query Match             15.1%;  Score 68.5;  DB 15;  Length 1943;
  Best Local Similarity   23.3%;  Pred. No. 51;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db       1019 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 1062

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       1063 DILNQMIEQAPLINKLISWLGPQDLS 1088


RESULT 10
US-09-895-913A-12
; Sequence 12, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
;  APPLICANT: Kleanthous, Harold
;  APPLICANT:  Al-Garawi, Amal
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tomb, Jean Francois
;  APPLICANT:  Oomen, Raymond P.
;  TITLE OF INVENTION: Identification of Polynucleotides
;  TITLE OF INVENTION:  Encoding No. US20020160456A1el Helicobacter Polypeptides in the
;  TITLE OF INVENTION:  Genome
;  FILE REFERENCE: 06132/043002
;  CURRENT APPLICATION NUMBER: US/09/895,913A
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: US 08/881,227
;  PRIOR FILING DATE: 1997-06-24
;  NUMBER OF SEQ ID NOS: 368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 1974
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-895-913A-12

  Query Match             15.1%;  Score 68.5;  DB 9;  Length 1974;
  Best Local Similarity   23.3%;  Pred. No. 52;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db       1050 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 1093

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       1094 DILNQMIEQAPLINKLISWLGPQDLS 1119


RESULT 11
US-10-335-977-5265
; Sequence 5265, Application US/10335977
; Publication No. US20040052799A1
;  GENERAL INFORMATION:
;       APPLICANT: DOUGLAS SMITH et al
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
;                           RELATING TO HELICOBACTER PYLORI FOR
;                           DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 10031
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: IBM PC Compatible
;            OPERATING SYSTEM: Windows NT 4.0
;            SOFTWARE: UNIX
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/335,977
;            FILING DATE: 30-Dec-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/993,002
;            FILING DATE: 17-DEC-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: GTN-018
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)742-4214
;  INFORMATION FOR SEQ ID NO: 5265:
;     SEQUENCE CHARACTERISTICS:
;            LENGTH: 1120 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;            ORGANISM: Helicobacter pylori
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION: (B) LOCATION 1...1120
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5265:
US-10-335-977-5265

  Query Match             14.9%;  Score 67.5;  DB 15;  Length 1120;
  Best Local Similarity   23.3%;  Pred. No. 34;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSK--RVVIEDDRID 64
              || : || :                  |||   | |:   |:|    :|  :   ::
Db        833 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQSFLDKTTFG 876

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
```

```
              |:|  | ::|||::       |: :|
Db        877 DILNQMINQAPLINKLISWLGPQDLS 902


RESULT 12
US-10-335-977-5266
; Sequence 5266, Application US/10335977
; Publication No. US20040052799A1
;  GENERAL INFORMATION:
;       APPLICANT: DOUGLAS SMITH et al
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
;                           RELATING TO HELICOBACTER PYLORI FOR
;                           DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 10031
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: IBM PC Compatible
;            OPERATING SYSTEM: Windows NT 4.0
;            SOFTWARE: UNIX
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/335,977
;            FILING DATE: 30-Dec-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/993,002
;            FILING DATE: 17-DEC-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: GTN-018
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)742-4214
;  INFORMATION FOR SEQ ID NO: 5266:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1576 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;            ORGANISM: Helicobacter pylori
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION:  (B) LOCATION 1...1576
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5266:
US-10-335-977-5266

  Query Match             14.9%;  Score 67.5;  DB 15;  Length 1576;
  Best Local Similarity   23.3%;  Pred. No. 52;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSK--RVVIEDDRID 64
              || : || :                  |||   | |:   |:|    :|  :  ::
Db       1289 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQSFLDKTTFG 1332

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       1333 DILNQMINQAPLINKLISWLGPQDLS 1358


RESULT 13
US-10-335-977-5267
; Sequence 5267, Application US/10335977
; Publication No. US20040052799A1
;  GENERAL INFORMATION:
;       APPLICANT: DOUGLAS SMITH et al
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
;                           RELATING TO HELICOBACTER PYLORI FOR
;                           DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 10031
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: IBM PC Compatible
;            OPERATING SYSTEM: Windows NT 4.0
;            SOFTWARE: UNIX
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/335,977
;            FILING DATE: 30-Dec-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/993,002
;            FILING DATE: 17-DEC-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: GTN-018
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)742-4214
;  INFORMATION FOR SEQ ID NO: 5267:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1775 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;            ORGANISM: Helicobacter pylori
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION:  (B) LOCATION 1...1775
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5267:
US-10-335-977-5267

  Query Match             14.9%;  Score 67.5;  DB 15;  Length 1775;
  Best Local Similarity   23.3%;  Pred. No. 61;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSK--RVVIEDDRID 64
              || : || :                  |||   | |:   |:|    :|  :  ::
Db       1488 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQSFLDKTTFG 1531

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       1532 DILNQMINQAPLINKLISWLGPQDLS 1557


RESULT 14
US-10-282-122A-59046
; Sequence 59046, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
```

```
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59046
;   LENGTH: 3194
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-10-282-122A-59046

  Query Match             14.9%;  Score 67.5;  DB 15;  Length 3194;
  Best Local Similarity   23.3%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSK--RVVIEDDRID 64
              || : || :                |||   | |:   |:|    :|  :  ::
Db       2265 YGKQSLSSF----------------ANNFVPGGLNQALDKIGSSSDAKDLQSFLDKTTFG 2308

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       2309 DILNQMINQAPLINKLISWLGPQDLS 2334


RESULT 15
US-10-369-493-3737
; Sequence 3737, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3737
;   LENGTH: 932
;   TYPE: PRT
;   ORGANISM: Neurospora crassa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(932)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3737

  Query Match             14.8%;  Score 67;  DB 15;  Length 932;
  Best Local Similarity   25.7%;  Pred. No. 31;
  Matches   19;  Conservative    8;  Mismatches   25;  Indels   22;  Gaps    2;

Qy         31 DTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDD----------------------VLKN 69
              | |   |:  |   |:    |:   | :::  | |                      |||
Db        448 DRNWTLGSASNSNHPE-SSAGQQDNVAVQEPEIQDSAVDLDRGYLSGTDADPRQRNVLKK 506

Qy         70 MTDKAPLVSNSPKT 83
                ||| :: ||||
Db        507 RESKAPALAESPKT 520


Search completed: March 28, 2005, 13:30:18
Job time : 1200 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 28, 2005, 10:45:49 ; Search time 50 Seconds
(without alignments)
163.568 Million cell updates/sec

Title: US-10-071-645-2
Perfect score: 453
Sequence: 1 MWEVLPYGDEKLSPYGDGGD.....VLKNMTDKAPLVSNSPKTMS 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	15.7	403	2	T09634	phosphoglycerate k
2	68.5	15.1	1943	2	B64596	toxin-like outer m
3	67.5	14.9	3194	2	D71917	toxin-like outer m
4	67.5	14.9	4717	2	T41581	hypothetical coile
5	65.5	14.5	579	2	S37953	hypothetical prote
6	65	14.3	959	2	B44402	nuclear pore compl
7	64.5	14.2	170	2	T09601	DnaJ protein homol
8	64.5	14.2	199	2	A96910	pemK family of DNA
9	64	14.1	1142	2	S36845	myosin-binding pro
10	63.5	14.0	1445	2	T14913	CAT8 protein - yea
11	63	13.9	138	2	A90258	conserved hypothet
12	63	13.9	152	1	A24427	interleukin-3 prec
13	63	13.9	152	1	S42720	interleukin-3 prec
14	63	13.9	529	2	T10612	hypothetical prote
15	63	13.9	614	2	T10862	phaseolin G-box bi
16	63	13.9	934	2	T39941	hypothetical prote
17	63	13.9	4589	2	T14914	dynein beta heavy
18	62.5	13.8	423	2	T09338	DnaJ-like protein
19	62.5	13.8	458	2	C85436	hypothetical prote
20	62.5	13.8	574	2	T01131	hypothetical prote
21	62.5	13.8	1371	2	T29019	hypothetical prote
22	62	13.7	315	2	S33362	L-lactate dehydrog
23	62	13.7	324	2	D37874	pectic enzyme secr
24	62	13.7	354	2	I47021	pectic enzyme secr
25	61.5	13.6	388	2	H64427	hypothetical prote
26	61.5	13.6	447	2	F64706	UDP-MurNac-tripept
27	61.5	13.6	527	2	T37842	probable transcrip
28	61.5	13.6	710	2	S40934	hypothetical prote
29	61	13.5	416	2	T40872	probable cell wall
30	61	13.5	778	2	B86218	protein T27G7.20 [
31	60.5	13.4	163	2	H83331	anthranilate dioxy
32	60.5	13.4	218	2	AG3554	multiple sugar-bin
33	60.5	13.4	2215	2	T16871	hypothetical prote
34	60	13.2	365	2	D70322	probable phosphori
35	60	13.2	396	2	A36339	FLO protein - gard
36	60	13.2	553	1	B55483	transcription init
37	60	13.2	726	2	S22258	probable protein k
38	60	13.2	1444	1	A30588	140K adhesin precu
39	60	13.2	1608	2	T13216	minor capsid prote
40	59.5	13.1	393	1	S22520	myb-related protei
41	59.5	13.1	447	2	D71812	udp-n-acetylmuramy
42	59.5	13.1	674	2	S74506	ribonuclease E - S
43	59	13.0	278	2	F97356	hypothetical secre
44	59	13.0	346	2	G83127	conserved hypothet
45	59	13.0	412	2	AB0578	hypothetical zinc-

ALIGNMENTS

RESULT 1
T09634
phosphoglycerate kinase (EC 2.7.2.3) - Lactobacillus delbrueckii
C;Species: Lactobacillus delbrueckii
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T09634
R;Branny, P.; Delatorre, F.; Garel, J.R.
Microbiology 144, 905-914, 1998
A;Title: An operon coding for three glycolytic enzymes in Lactobacillus delbrueckii subs
A;Reference number: Z16788; MUID:98240227; PMID:9579064
A;Accession: T09634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-403 <BRA>
A;Cross-references: UNIPROT:O32756; EMBL:AJ000339; NID:g2624189; PIDN:CAA04015.1; PID:g2
A;Experimental source: subsp. bulgaricus, strain B107
C;Genetics:
A;Gene: pgk
C;Superfamily: phosphoglycerate kinase
C;Keywords: phosphotransferase

Query Match 15.7%; Score 71; DB 2; Length 403;
Best Local Similarity 39.5%; Pred. No. 4.3;
Matches 17; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

```
Qy         19 GDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDD 61
              |||  : : | || :| ||  ::|  ||||:   |    |  :|
Db        110 GDVVVLENTRFQDIDNDFGKRESKNDPKLGEYWASLGDVFVND 152
```

RESULT 2
B64596
toxin-like outer membrane protein HP0610 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: B64596
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: B64596
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1943 <TOM>
A;Cross-references: UNIPROT:O25331; GB:AE000575; GB:AE000511; NID:g2313730; PIDN:AAD0767
C;Genetics:
A;Start codon: GTG

```
  Query Match             15.1%;  Score 68.5;  DB 2;  Length 1943;
  Best Local Similarity   23.3%;  Pred. No. 47;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRID 64
              || : || :                  |||   | |:   |:|    :| :   ::
Db       1019 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQNFLDKTTFG 1062

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       1063 DILNQMIEQAPLINKLISWLGPQDLS 1088


RESULT 3
D71917
toxin-like outer membrane protein jhp0556 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: D71917
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: D71917
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3194 <ARN>
A;Cross-references: UNIPROT:Q9ZLM3; GB:AE001488; GB:AE001439; NID:g4155100; PIDN:AAD0613
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0556

  Query Match             14.9%;  Score 67.5;  DB 2;  Length 3194;
  Best Local Similarity   23.3%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative   15;  Mismatches   28;  Indels   23;  Gaps    3;

Qy          7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSK--RVVIEDDRID 64
              || : || :                  |||   | |:   |:|    :|  :  ::
Db       2265 YGKQSLSSF---------------ANNFVPGGLNQALDKIGSSSDAKDLQSFLDKTTFG 2308

Qy         65 DVLKNMTDKAPLVSN-----SPKTMS 85
              |:|  | ::|||::       |: :|
Db       2309 DILNQMINQAPLINKLISWLGPQDLS 2334


RESULT 4
T41581
hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41581
R;Murphy, L.; Harris, D.; Wood, V.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z22002
A;Accession: T41581
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4717 <MUR>
A;Cross-references: UNIPROT:O94248; EMBL:AL031546; PIDN:CAA20864.1; GSPDB:GN00068; SPDB:
A;Experimental source: strain 972h(-)
C;Genetics:
A;Gene: SPDB:SPCC737.08
A;Map position: 3
C;Superfamily: midasin (AAA ATPase with von Willebrand factor type A (vWA) domain)

  Query Match             14.9%;  Score 67.5;  DB 2;  Length 4717;
  Best Local Similarity   27.8%;  Pred. No. 1.6e+02;
  Matches   22;  Conservative   15;  Mismatches   23;  Indels   19;  Gaps    4;

Qy          8 GDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRV--VIEDDRIDD 65
              ||::|    || |:  |              |::  | : | | :|::|  : ||: |:|
Db       4364 GDKQLQ---DGQDISDIKQ----------TGEDTLPTEFGSINQSEKVFELSEDEDIED 4409

Qy         66 VLKNMTDKAPLVSNSPKTM 84
                | :    |   ::| |  |
Db       4410 ELPDYNVK---ITNLPAAM 4425


RESULT 5
S37953
hypothetical protein YKL124w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YKL529
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: S37953; S37952; S25361
R;Ramezani Rad, M.; Xu, G.; Kirchrath, L.; Fritz, C.; Keuchel, H.; Hollenberg, C.P.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37953
A;Accession: S37953
A;Molecule type: DNA
A;Residues: 1-120 <RAM>
A;Cross-references: UNIPROT:P32343; EMBL:Z28123; MIPS:YKL124w
A;Experimental source: strain S288C
R;Jacquier, A.; Legrain, P.; Colleaux, L.; Richard, G.F.; Thierry, A.; Dujon, B.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37938
A;Accession: S37952
A;Molecule type: DNA
A;Residues: 41-579 <JAC>
A;Cross-references: EMBL:Z28123; MIPS:YKL124w
A;Experimental source: strain S288C
R;Colleaux, L.; Richard, G.F.; Thierry, A.; Dujon, B.
Yeast 8, 325-336, 1992
A;Title: Sequence of a segment of yeast chromosome XI identifies a new mitochondrial car
A;Reference number: S25357; MUID:92383952; PMID:1514329
A;Accession: S25361
A;Molecule type: DNA
A;Residues: 41-579 <COL>
A;Cross-references: EMBL:S44213; NID:g254447; PIDN:AAB23074.1; PID:g254452
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SSH4
A;Cross-references: SGD:S0001607; MIPS:YKL124w
A;Map position: 11L

  Query Match             14.5%;  Score 65.5;  DB 2;  Length 579;
  Best Local Similarity   22.3%;  Pred. No. 26;
  Matches   29;  Conservative   16;  Mismatches   32;  Indels   53;  Gaps    5;

Qy          7 YGDEKLSPYGDGGDV--------GQIFSCRLQDTNNFFGA-----------GQNKRPPK 46
              ||   : |  :| ::         |:    |  ||:||||:           |  :|
Db        420 YGQIGIPPAYNGTEIKKDTILQKGEELPPRYADTDNFFGSMKVKEGSSSRITAQTSKP-- 477

Qy         47 LGQIGRSKRVVIEDDRIDDVL-------KNMTD-----------------------KAP 75
              |  :|  :|:    || ::|         | ||                       | |
Db        478 LWSVGTYERISSNFDRENNVYHDSLETDDNNTDNNVNNNDENAGCNENSPLLEDDGNKRP 537

Qy         76 LVSNSPKTMS 85
                ||:|: :|
Db        538 ENSNTPREVS 547


RESULT 6
B44402
nuclear pore complex protein NUP100 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKL068w; protein YKL336
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B44402; S37890; S44518; S28536; S39173
R;Wente, S.R.; Rout, M.P.; Blobel, G.
J. Cell Biol. 119, 705-723, 1992
A;Title: A new family of yeast nuclear pore complex proteins.
A;Reference number: A44402; MUID:93054906; PMID:1385442
A;Accession: B44402
```

A;Molecule type: DNA
A;Residues: 1-959 <WEN>
A;Cross-references: UNIPROT:Q02629; EMBL:Z15035; NID:g4073; PIDN:CAA78753.1; PID:g4074
A;Note: sequence extracted from NCBI backbone (NCBIP:117135)
R;Rasmussen, S.; von Wettstein, D.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37872
A;Accession: S37890
A;Molecule type: DNA
A;Residues: 1-959 <RAS>
A;Cross-references: EMBL:Z28068; NID:g486094; PID:g486095; MIPS:YKL068w
A;Experimental source: strain S288C
R;Rasmussen, S.W.
Yeast 10, 69-74, 1994
A;Title: Sequence of a 20.7 kb region of yeast chromosome XI includes the NUP100 gene, a
rp in addition to seven ORFs with weak or no significant similarity to known proteins.
A;Reference number: S44513
A;Accession: S44518
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-959 <RA2>
A;Cross-references: EMBL:X75780; NID:g433625; PID:g433631
C;Genetics:
A;Gene: SGD:NUP100
A;Cross-references: SGD:S0001551; MIPS:YKL068w
A;Map position: 11L
C;Keywords: nucleus

  Query Match 14.3%; Score 65; DB 2; Length 959;
  Best Local Similarity 34.0%; Pred. No. 53;
  Matches 16; Conservative 9; Mismatches 14; Indels 8; Gaps 2;

Qy 7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRS 53
Db 204 FGNQQSSAFGTNNQQGSLFGQQSQNTNNAFG-NQN-------QLGGS 242

RESULT 7
T09601
DnaJ protein homolog - alfalfa (fragment)
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T09601
R;Frugis, G.
submitted to the EMBL Data Library, April 1996
A;Description: Isolation of an alfalfa DnaJ-like gene and a Y13 proteasome subunit homol
A;Reference number: Z16763
A;Accession: T09601
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-170 <FRU>
A;Cross-references: UNIPROT:O24075; EMBL:Z71997
A;Experimental source: variety Rangelander; non-embryogenic callus
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
C;Keywords: molecular chaperone

  Query Match 14.2%; Score 64.5; DB 2; Length 170;
  Best Local Similarity 27.6%; Pred. No. 8.6;
  Matches 21; Conservative 10; Mismatches 42; Indels 3; Gaps 2;

Qy 7 YGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRI--D 64
Db 12 YGEDALKE-GMGGGAGSSFHNPFDILQSFFGAGLGGGGPSRARRQKQGEDVVHSIKVSLE 70

Qy 65 DVLKNMTDKAPLVSNS 80
Db 71 DVYNGTTKKLSLSRNA 86

RESULT 8
A96910
pemK family of DNA-binding proteins [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: A96910
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A96910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-199 <KUR>
A;Cross-references: UNIPROT:Q97MV8; GB:AE001437; PIDN:AAK78068.1; PID:g15022906; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0083

  Query Match 14.2%; Score 64.5; DB 2; Length 199;
  Best Local Similarity 23.9%; Pred. No. 10;
  Matches 22; Conservative 22; Mismatches 27; Indels 21; Gaps 4;

Qy 1 MWEVLPYGDEKL-------SPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRS 53
Db 59 VYEYIKWANDKIAINNNVESSYGTIPKRGEIWTCQLGEN---IGSEENKIRP-------- 107

Qy 54 KRVVIEDDRIDDVLKNMTDKAPLVSNSPKTMS 85
Db 108 -AIIIQNDTGNE--KGPTTIIVPISNRPKKIS 136

RESULT 9
S36845
myosin-binding protein C, fast-type muscle - human
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S36845
R;Weber, F.E.; Vaughan, K.T.; Reinach, F.C.; Fischman, D.A.
Eur. J. Biochem. 216, 661-669, 1993
A;Title: Complete sequence of human fast-type and slow-type muscle myosin-binding-protei
A;Reference number: S36845; MUID:93387319; PMID:8375400
A;Accession: S36845
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1142 <WEB>
A;Cross-references: UNIPROT:Q14324; EMBL:X73113; NID:g402646; PIDN:CAA51544.1; PID:g4026

  Query Match 14.1%; Score 64; DB 2; Length 1142;
  Best Local Similarity 31.2%; Pred. No. 83;
  Matches 15; Conservative 10; Mismatches 17; Indels 6; Gaps 2;

Qy 24 IFSCRLQD---TNNFFGAGQNKRPPK---LGQIGRSKRVVIEDDRIDD 65
Db 457 VFKCEVSDEKVTGKWYKNGVEVRPSKRITISHVGRFHKLVIDDVRPED 504

RESULT 10
T14913
CAT8 protein - yeast (Kluyveromyces marxianus var. lactis)
C;Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
C;Accession: T14913
R;Georis, I.; Breunig, K.D.; Vandenhaute, J.
submitted to the EMBL Data Library, June 1998
A;Description: Sequence of KlCAT8.
A;Reference number: Z18263
A;Accession: T14913
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1445 <GEO>
A;Cross-references: UNIPROT:O74229; EMBL:AF070974; NID:g3228690; PID:g3228691; PIDN:AAC2
C;Genetics:
A;Gene: CAT8
C;Superfamily: GAL4 zinc binuclear cluster homology
F;178-215/Domain: GAL4 zinc binuclear cluster homology <GL4>

```
  Query Match             14.0%;  Score 63.5;  DB 2;  Length 1445;
  Best Local Similarity   33.9%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    8;  Mismatches   20;  Indels    9;  Gaps    2;

Qy         38 AGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMT--------DKAPLVSNSPKTMS 85
              | : || ||  :   ||:| :|||   | | |:|           || : || ::
Db        999 AVKRKRDPKATE-NTSKKVKVEDDHSQDNLHNITTGETTDTVHSEELVKDVPKELN 1053


RESULT 11
A90258
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: A90258
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A90258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-138 <KUR>
A;Cross-references: UNIPROT:Q97Z73; GB:AE006641; NID:g13814244; PIDN:AAK41320.1; GSPDB:G
C;Genetics:
A;Gene: SSO1056

  Query Match             13.9%;  Score 63;  DB 2;  Length 138;
  Best Local Similarity   32.6%;  Pred. No. 10;
  Matches   15;  Conservative    8;  Mismatches   21;  Indels    2;  Gaps    1;

Qy         31 DTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPL 76
              ||  |   |  : || | ::     |  :| || |:||   :: |:
Db         77 DTKGF--QGVEEVPPGLLEVSYELIVYSKDSRISDILKEAEERCPM 120


RESULT 12
A24427
interleukin-3 precursor - human
N;Alternate names: hematopoietic growth factor; IL-3; mast-cell growth factor; multi-CSF
C;Species: Homo sapiens (man)
C;Date: 05-Jun-1988 #sequence_revision 15-Nov-1996 #text_change 09-Jul-2004
C;Accession: A24427; A27613; A27413; A60455; A30615
R;Yang, Y.C.; Ciarletta, A.B.; Temple, P.A.; Chung, M.P.; Kovacic, S.; Witek-Giannotti,
Cell 47, 3-10, 1986
A;Title: Human IL-3 (multi-CSF): identification by expression cloning of a novel hematop
A;Reference number: A90885; MUID:87002490; PMID:3489530
A;Accession: A24427
A;Molecule type: DNA
A;Residues: 1-26,'S',28-152 <YAN>
A;Cross-references: UNIPROT:P08700; GB:M14743; NID:g186326; PIDN:AAA59146.1; PID:g307059
R;Otsuka, T.; Miyajima, A.; Brown, N.; Otsu, K.; Abrams, J.; Saeland, S.; Caux, C.; De W
.; Yokota, T.
J. Immunol. 140, 2288-2295, 1988
A;Title: Isolation and characterization of an expressible cDNA encoding human IL-3. Indu
A;Reference number: A27613; MUID:88170808; PMID:3127463
A;Accession: A27613
A;Molecule type: mRNA
A;Residues: 1-152 <OTS>
A;Cross-references: GB:M20137; NID:g186328; PIDN:AAA59147.1; PID:g307060
R;Dorssers, L.; Burger, H.; Bot, F.; Delwel, R.; Geurts van Kessel, A.H.M.; Lowenberg, B
Gene 55, 115-124, 1987
A;Title: Characterization of a human multilineage-colony-stimulating factor cDNA clone i
A;Reference number: A27413; MUID:87305582; PMID:3497843
A;Accession: A27413
A;Molecule type: mRNA
A;Residues: 1-152 <DOR>
A;Cross-references: GB:M17115; NID:g181151; PIDN:AAA35725.1; PID:g181152
R;Urdal, D.L.; Price, V.; Sassenfeld, H.M.; Cosman, D.; Gillis, S.; Park, L.S.
Ann. N. Y. Acad. Sci. 554, 167-176, 1989
A;Title: Molecular characterization of colony-stimulating factors and their receptors: h
A;Reference number: A60455; MUID:89285823; PMID:2544122
A;Accession: A60455
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 20-26,'S',28-152 <URD>
A;Note: the authors studied the polymorphism of this protein and found 27-Pro to be some
R;Yang, L.H.; Chen, C.Q.; Zhou, B.; Chen, N.C.; Yao, L.B.; Su, C.Z.
Acta Biochim. Biophys. Sin. 26, 477-484, 1994
A;Title: High-level expression of the gene encoding human interleukin-3 in Escherichia c
A;Reference number: A30615
A;Accession: A30615
A;Molecule type: mRNA
A;Residues: 'M',20-128,'R',130-147,'R',149-152 <YA2>
C;Comment: This glycoprotein, produced by bone marrow cells, by mitogen or antigen-activ
erentiation of hematopoietic cells.
C;Genetics:
A;Gene: GDB:IL3
A;Cross-references: GDB:120095; OMIM:147740
A;Map position: 5q23-5q31
A;Introns: 54/3; 68/3; 98/3; 112/3
C;Superfamily: interleukin-3
C;Keywords: cytokine; glycoprotein; growth factor; lymphokine; mitogen; monomer; T-cell
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-152/Product: interleukin-3 #status predicted <MAT>
F;34,89/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;35-103/Disulfide bonds: #status predicted

  Query Match             13.9%;  Score 63;  DB 1;  Length 152;
  Best Local Similarity   30.5%;  Pred. No. 11;
  Matches   18;  Conservative    8;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  |  |       | | | |    |: : :     |: :|||:    || : :|
Db         53 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 111


RESULT 13
S42720
interleukin-3 precursor - chimpanzee
N;Alternate names: hematopoietic growth factor; IL-3; mast-cell growth factor; multi-CSF
C;Species: Pan troglodytes (chimpanzee)
C;Date: 13-Jan-1995 #sequence_revision 15-Nov-1996 #text_change 09-Jul-2004
C;Accession: S42720
R;Burger, H.; Mostert, M.C.; Kok, E.M.; Wagemaker, G.; Dorssers, L.C.J.
Biochim. Biophys. Acta 1217, 195-198, 1994
A;Title: Cloning and expression of interleukin-3 genes of chimpanzee and New World monke
A;Reference number: S42720; MUID:94153996; PMID:8110834
A;Accession: S42720
A;Molecule type: DNA
A;Residues: 1-152 <BUR>
A;Cross-references: UNIPROT:Q28809; EMBL:X74875; NID:g459542; PIDN:CAA52863.1; PID:g8291
C;Comment: This glycoprotein, produced by bone marrow cells, by mitogen or antigen-activ
erentiation of hematopoietic cells.
C;Genetics:
A;Introns: 54/3; 68/3; 98/3; 112/3
C;Superfamily: interleukin-3
C;Keywords: cytokine; glycoprotein; growth factor; lymphokine; mitogen; monomer; T-cell
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-152/Product: interleukin-3 #status predicted <MAT>
F;34,89/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;35-103/Disulfide bonds: #status predicted

  Query Match             13.9%;  Score 63;  DB 1;  Length 152;
  Best Local Similarity   30.5%;  Pred. No. 11;
  Matches   18;  Conservative    8;  Mismatches   27;  Indels    6;  Gaps    1;

Qy         29 LQDTNNFFGAGQ------NKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSP 81
              | | ||  |  |       | | | |    |: : :     |: :|||:    || : :|
Db         53 LLDFNNLNGEDQDILMENNLRRPNLEAFNRAVKSLQNASAIESILKNLLPCLPLATAAP 111


RESULT 14
```

```
T10612
hypothetical protein F21C20.50 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10612
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16991
A;Accession: T10612
A;Molecule type: DNA
A;Residues: 1-529 <BEV>
A;Cross-references: UNIPROT:Q9SVH7; EMBL:AL080254; GSPDB:GN00062; ATSP:F21C20.50
A;Experimental source: cultivar Columbia; BAC clone F21C20
C;Genetics:
A;Gene: ATSP:F21C20.50
A;Map position: 4
A;Introns: 23/2; 89/1

  Query Match             13.9%;  Score 63;  DB 2;  Length 529;
  Best Local Similarity   31.9%;  Pred. No. 45;
  Matches   22;  Conservative   10;  Mismatches   25;  Indels   12;  Gaps    3;

Qy         16 GDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDD----RIDDVLKNMT 71
              |||    :     ||:::  |||   :  |::|:   |||  : |      | ||:
Db        175 GDGQVARRELVLGLQNSSGTFGA--RRVAPEVGRADASKRPRVSDHRWGFRYDDM----- 227

Qy         72 DKAPLVSNS 80
                ||  |||
Db        228 -AAPFTSNS 235


RESULT 15
T10862
phaseolin G-box binding protein PG2 - kidney bean (fragment)
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10862
R;Kawagoe, Y.; Murai, N.
Plant Physiol. 117, 720, 1998
A;Title: PEP2, a second bHLH protein binding to the E-box motif of the bean seed-storage
A;Reference number: Z17191
A;Accession: T10862
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-614 <KAW>
A;Cross-references: UNIPROT:Q41102; EMBL:U18349; NID:g1142620; PID:g1142621
A;Experimental source: strain Contender
C;Genetics:
A;Gene: PG2
C;Function:
A;Description: transcription factor

  Query Match             13.9%;  Score 63;  DB 2;  Length 614;
  Best Local Similarity   22.7%;  Pred. No. 53;
  Matches   22;  Conservative   15;  Mismatches   30;  Indels   30;  Gaps    3;

Qy          3 EVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDR 62
              |:| : | | : || ||  | : :      :|||             |:|  | : |:
Db        309 EILSFSDSKRTSYGGGGGGGV-NGNSNSNSNFFS-------------GQSPFVAVADEN 354

Qy         63 ----------------IDDVLKNMTDKAPLVSNSPKT 83
                             || : : | :| | : : |:
Db        355 NNNNNGKRRSPNSRGSNDDGMLSFTSRAILPATNLKS 391


Search completed: March 28, 2005, 13:02:30
Job time : 54 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 28, 2005, 06:42:28 ; Search time 131 Seconds
(without alignments)
332.265 Million cell updates/sec

Title: US-10-071-645-2
Perfect score: 453
Sequence: 1 MWEVLPYGDEKLSPYGDGGD.....VLKNMTDKAPLVSNSPKTMS 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	391	86.3	78	2	Q7Z7J9	Q7z7j9 homo sapien
2	380	83.9	78	2	Q6QWF9	Q6qwf9 mus musculu
3	380	83.9	78	2	Q9JI15	Q9ji15 rattus norv
4	291.5	64.3	79	2	Q6P6Z4	Q6p6z4 xenopus lae
5	270.5	59.7	79	2	Q6DHN6	Q6dhn6 brachydanio
6	259.5	57.3	79	2	Q78WH7	Q78wh7 mus musculu
7	259.5	57.3	79	2	Q9Z2N6	Q9z2n6 rattus norv
8	253.5	56.0	79	2	Q96S95	Q96s95 homo sapien
9	77.5	17.1	246	2	Q88EZ7	Q88ez7 pseudomonas
10	71.5	15.8	215	2	Q8CPA2	Q8cpa2 staphylococ
11	71	15.7	403	1	PGK_LACDE	O32756 lactobacill
12	71	15.7	403	1	PGK_LACDL	Q8giz5 lactobacill
13	71	15.7	403	1	PGK_LACJO	P62413 lactobacill
14	71	15.7	620	2	Q8QQC0	Q8qqc0 equus cabal
15	71	15.7	620	2	Q8BE74	Q8be74 equinus pap
16	70.5	15.6	445	2	Q74IF4	Q74if4 lactobacill
17	70.5	15.6	737	2	Q654P9	Q654p9 oryza sativ
18	70	15.5	279	1	NADE_SULTO	Q96yl5 sulfolobus
19	68.5	15.1	1943	2	O25331	O25331 helicobacte
20	68	15.0	186	2	Q8MUK9	Q8muk9 strongyloce
21	68	15.0	485	2	Q7UQA3	Q7uqa3 rhodopirell
22	68	15.0	505	2	Q8D3D8	Q8d3d8 wiggleswort
23	67.5	14.9	3194	2	Q9ZLM3	Q9zlm3 helicobacte
24	67.5	14.9	4717	2	O94248	O94248 schizosacch
25	67	14.8	897	2	Q7S7I3	Q7s7i3 neurospora
26	67	14.8	945	2	Q8X0U9	Q8x0u9 neurospora
27	67	14.8	1464	2	Q871F7	Q871f7 neurospora
28	66.5	14.7	251	2	Q7N5M9	Q7n5m9 photorhabdu
29	66.5	14.7	378	2	Q6MPH9	Q6mph9 bdellovibri
30	66.5	14.7	757	2	Q66VR6	Q66vr6 micropogoni
31	66	14.6	263	2	Q8N0C6	Q8n0c6 daphnia spi
32	66	14.6	272	2	Q8N0D0	Q8n0d0 daphnia spi
33	65.5	14.5	298	2	Q65WD4	Q65wd4 mannheimia
34	65.5	14.5	560	2	Q8PFN6	Q8pfn6 xanthomonas
35	65.5	14.5	579	1	YKM4_YEAST	P32343 saccharomyc
36	65.5	14.5	894	2	Q6FE66	Q6fe66 acinetobact
37	65.5	14.5	1164	2	Q8PT00	Q8pt00 methanosarc
38	65	14.3	427	2	Q8ZU72	Q8zu72 pyrobaculum
39	65	14.3	568	2	Q7R1Q1	Q7r1q1 giardia lam
40	65	14.3	959	1	N100_YEAST	Q02629 saccharomyc
41	65	14.3	1190	2	Q782R4	Q782r4 gallid herp
42	65	14.3	1190	2	Q9WSY3	Q9wsy3 marek's dis
43	64.5	14.2	170	2	O24075	O24075 medicago sa
44	64.5	14.2	199	2	Q97MV8	Q97mv8 clostridium
45	64.5	14.2	391	2	Q83VI7	Q83vi7 sphingomona

ALIGNMENTS

RESULT 1
Q7Z7J9
ID Q7Z7J9 PRELIMINARY; PRT; 78 AA.
AC Q7Z7J9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase II inhibitor alpha.
GN Name=CaMKIINalpha;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li N., Cao X.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY204901; AAO49802.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
KW Kinase.
SQ SEQUENCE 78 AA; 8553 MW; 31866C5DA10B606F CRC64;

Query Match 86.3%; Score 391; DB 2; Length 78;
Best Local Similarity 98.7%; Pred. No. 3.4e-37;
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAP 75
              ||||||||||||||||
Db         61 DRIDDVLKNMTDKAP 75
```

RESULT 2
Q6QWF9
ID Q6QWF9 PRELIMINARY; PRT; 78 AA.
AC Q6QWF9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CaMKII inhibitor protein alpha.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss;
RA Sougata S., Rangarajan P.N.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY523601; AAS02090.1; -.
SQ SEQUENCE 78 AA; 8513 MW; A89A9E5DA11B8E61 CRC64;

```
  Query Match             83.9%;  Score 380;  DB 2;  Length 78;
  Best Local Similarity   96.0%;  Pred. No. 6.3e-36;
  Matches   72;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              | ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db          1 MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQSKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAP 75
              |||||||| ||||||
Db         61 DRIDDVLKTMTDKAP 75


RESULT 3
Q9JI15
ID   Q9JI15      PRELIMINARY;      PRT;    78 AA.
AC   Q9JI15;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CaM-kinase II inhibitor alpha.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21106800; PubMed=11182241; DOI=10.1016/S0306-4522(00)00520-0;
RA   Chang B.H., Mukherji S., Soderling T.R.;
RT   "Calcium/calmodulin-dependent protein kinase II inhibitor protein:
RT   localization of isoforms in rat brain.";
RL   Neuroscience 102:767-777(2001).
DR   EMBL; AF271156; AAF75281.1; -.
DR   GO; GO:0016301; F:kinase activity; IEA.
KW   Kinase.
SQ   SEQUENCE   78 AA;  8513 MW;  A89A9E5DA11B8E61 CRC64;

  Query Match             83.9%;  Score 380;  DB 2;  Length 78;
  Best Local Similarity   96.0%;  Pred. No. 6.3e-36;
  Matches   72;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIED 60
              | ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db          1 MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQSKRPPKLGQIGRSKRVVIED 60

Qy         61 DRIDDVLKNMTDKAP 75
              |||||||| ||||||
Db         61 DRIDDVLKTMTDKAP 75


RESULT 4
Q6P6Z4
ID   Q6P6Z4      PRELIMINARY;      PRT;    79 AA.
AC   Q6P6Z4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC061942; AAH61942.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   79 AA;  8660 MW;  0F57CAE86FE750C2 CRC64;

  Query Match             64.3%;  Score 291.5;  DB 2;  Length 79;
  Best Local Similarity   72.4%;  Pred. No. 9.9e-26;
  Matches   55;  Conservative   10;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQI-FSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIE 59
              | |:||| |||:  ||  |:|||| ||||||||::|||  | ||||||||||:||||||
Db          1 MSEILPYSDEKMGHYGSDGEVGQISFSCRLQDTSSFFGGNQQKRPPKLGQIGRAKRVVIE 60

Qy         60 DDRIDDVLKNMTDKAP 75
              |||||:||| ::||:|
Db         61 DDRIDEVLKGVSDKSP 76


RESULT 5
Q6DHN6
ID   Q6DHN6      PRELIMINARY;      PRT;    79 AA.
AC   Q6DHN6;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Zgc:92198.
GN   Name=zgc:92198;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RA   Strausberg R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC075931; AAH75931.1; -.
SQ   SEQUENCE   79 AA;  8666 MW;  1210C2F714CA4896 CRC64;

  Query Match             59.7%;  Score 270.5;  DB 2;  Length 79;
  Best Local Similarity   70.7%;  Pred. No. 2.6e-23;
  Matches   53;  Conservative    9;  Mismatches   12;  Indels    1;  Gaps    1;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQI-FSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIE 59
              | ||||| : |:: ||   :| || |||||||||:||| |:||||||||||:| ||||
Db          1 MSEVLPYNEGKMNGYGADSEVSQISFSCRLQDTNSFFGTSQSKRPPKLGQIGRAKHVVIE 60

Qy         60 DDRIDDVLKNMTDKA 74
              |||||:||| ||||:
Db         61 DDRIDEVLKGMTDKS 75


RESULT 6
Q78WH7
ID   Q78WH7      PRELIMINARY;      PRT;    79 AA.
AC   Q78WH7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched
DE   library, clone:2900075A18 product:CAM-KII INHIBITORY PROTEIN
DE   homolog.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hippocampus;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK013788; BAB28995.1; -.
SQ   SEQUENCE   79 AA;  8628 MW;  B34820F5D6417CF3 CRC64;

  Query Match             57.3%;  Score 259.5;  DB 2;  Length 79;
  Best Local Similarity   65.4%;  Pred. No. 4.8e-22;
  Matches   51;  Conservative   10;  Mismatches   12;  Indels    5;  Gaps    2;

Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:|||::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYGEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 7
Q9Z2N6
ID   Q9Z2N6      PRELIMINARY;      PRT;    79 AA.
AC   Q9Z2N6;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CaM-KII inhibitory protein.
GN   Name=2900075A18Rik;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chang B.H., Mukherji S., Soderling T.R.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF041854; AAD02202.1; -.
DR   MGD; MGI:1920297; 2900075A18Rik.
SQ   SEQUENCE   79 AA;  8628 MW;  B34820F5D6417CF3 CRC64;

  Query Match             57.3%;  Score 259.5;  DB 2;  Length 79;
  Best Local Similarity   65.4%;  Pred. No. 4.8e-22;
  Matches   51;  Conservative   10;  Mismatches   12;  Indels    5;  Gaps    2;
```

```
Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:||||::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYGEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 8
Q96S95
ID   Q96S95      PRELIMINARY;      PRT;    79 AA.
AC   Q96S95;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CaM-KII inhibitory protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Li N., Zhang J., Wan T., Zhang W., Cao X.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY037149; AAK67630.1; -.
SQ   SEQUENCE   79 AA;  8658 MW;  160C31A092417CEE CRC64;

  Query Match             56.0%;  Score 253.5;  DB 2;  Length 79;
  Best Local Similarity   64.1%;  Pred. No. 2.3e-21;
  Matches   50;  Conservative   10;  Mismatches   13;  Indels    5;  Gaps    2;

Qy          1 MWEVLPYGDEKLSPYG---DGGDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVV 57
              | |:||| ::|:  :|   :| |:   |||||||||:||   | ||||||||||:||||
Db          1 MSEILPYSEDKMGRFGADPEGSDLS--FSCRLQDTNSFFAGNQAKRPPKLGQIGRAKRVV 58

Qy         58 IEDDRIDDVLKNMTDKAP 75
              |||||||||| | :| |
Db         59 IEDDRIDDVLKGMGEKPP 76


RESULT 9
Q88EZ7
ID   Q88EZ7      PRELIMINARY;      PRT;   246 AA.
AC   Q88EZ7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Conserved domain protein.
GN   OrderedLocusNames=PP4303;
OS   Pseudomonas putida (strain KT2440).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=160488;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22423060; PubMed=12534463;
RA   Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA   Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA   Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA   Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA   Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA   Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA   Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA   Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative analysis of the
RT   metabolically versatile Pseudomonas putida KT2440.";
RL   Environ. Microbiol. 4:799-808(2002).
DR   EMBL; AE016790; AAN69883.1; -.
DR   TIGR; PP4303; -.
KW   Complete proteome.
SQ   SEQUENCE   246 AA;  26916 MW;  9BFF373D77BA077A CRC64;

  Query Match             17.1%;  Score 77.5;  DB 2;  Length 246;
  Best Local Similarity   34.1%;  Pred. No. 1.6;
  Matches   28;  Conservative    7;  Mismatches   26;  Indels   21;  Gaps    5;

Qy          7 YGDEKLSPYGDGGDVGQIFSC------RLQDTNNFFGAGQNKRPPKLGQIGRSKRVVI-- 58
              :  : |||:| || |      |       ||  |    | |  || |     ||::|| :
Db         62 FSPDMLSPFGKGG-VNTYSYCQGDPVNRLDPTG---GTGILKRLP-----GRARRVAVVS 112

Qy         59 ----EDDRIDDVLKNMTDKAPL 76
                  |||:  | ||   |  |
Db        113 PVKATDDRVSPVFKNGPGKYQL 134


RESULT 10
Q8CPA2
ID   Q8CPA2      PRELIMINARY;      PRT;   215 AA.
AC   Q8CPA2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Negative regulator PhoU.
GN   OrderedLocusNames=SE1066;
OS   Staphylococcus epidermidis.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1282;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 12228;
RX   PubMed=12950922;
RA   Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Pu G., Yang J.,
RA   Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA   Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT   "Genome-based analysis of virulence genes in a non-biofilm-forming
RT   Staphylococcus epidermidis strain (ATCC 12228).";
RL   Mol. Microbiol. 49:1577-1593(2003).
DR   EMBL; AE016747; AAO04663.1; -.
DR   InterPro; IPR008170; PhoU.
DR   Pfam; PF01895; PhoU; 2.
KW   Complete proteome.
SQ   SEQUENCE   215 AA;  24918 MW;  C32C45E971178180 CRC64;

  Query Match             15.8%;  Score 71.5;  DB 2;  Length 215;
  Best Local Similarity   31.1%;  Pred. No. 6.8;
  Matches   19;  Conservative   13;  Mismatches   20;  Indels    9;  Gaps    2;

Qy         21 VGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNS 80
              :|::    |:| ||   | :||  | :      | |:  |: |||:  |: : : |: |
Db        124 MGKLAMLMLEDLNN---AIRNKDLPLI------KEVIERDEDIDDLYVNIVNTSYLIDND 174

Qy         81 P 81
              |
Db        175 P 175


RESULT 11
PGK_LACDE
ID   PGK_LACDE      STANDARD;      PRT;   403 AA.
AC   O32756;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Phosphoglycerate kinase (EC 2.7.2.3).
GN   Name=pgk;
OS   Lactobacillus delbrueckii (subsp. bulgaricus).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1585;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B107;
RX   MEDLINE=98240227; PubMed=9579064;
```

```
RA   Branny P., Delatorre F., Garel J.R.;
RT   "An operon encoding three glycolytic enzymes in Lactobacillus
RT   delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
RT   dehydrogenase, phosphoglycerate kinase and triosephosphate
RT   isomerase.";
RL   Microbiology 144:905-914(1998).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-phospho-
CC       D-glyceroyl phosphate.
CC   -!- PATHWAY: Second phase of glycolysis; second step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ000339; CAA04015.1; -.
DR   PIR; T09634; T09634.
DR   HSSP; P36204; 1VPE.
DR   HAMAP; MF_00145; -; 1.
DR   InterPro; IPR001576; PGK.
DR   Pfam; PF00162; PGK; 1.
DR   PIRSF; PIRSF000724; Pgk; 1.
DR   PRINTS; PR00477; PHGLYCKINASE.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Glycolysis; Kinase; Transferase.
SQ   SEQUENCE   403 AA;  42712 MW;  D90FC78FBB9FE397 CRC64;

  Query Match             15.7%;  Score 71;  DB 1;  Length 403;
  Best Local Similarity   39.5%;  Pred. No. 16;
  Matches   17;  Conservative    7;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         19 GDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDD 61
              |||  : : | || :| ||  ::|  ||||:   |   | :|
Db        110 GDVVVLENTRFQDIDNDFGKRESKNDPKLGEYWASLGDVFVND 152


RESULT 12
PGK_LACDL
ID   PGK_LACDL      STANDARD;      PRT;   403 AA.
AC   Q8GIZ5;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Phosphoglycerate kinase (EC 2.7.2.3).
GN   Name=pgk;
OS   Lactobacillus delbrueckii (subsp. lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=29397;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 88;
RA   Bourniquel A.A., Mollet B.;
RT   "Purification and characterization of the 3-phosphoglycerate kinase
RT   from the thermophile Lactobacillus delbrueckii subsp. lactis.";
RL   Int. Dairy J. 12:723-728(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-phospho-
CC       D-glyceroyl phosphate.
CC   -!- PATHWAY: Second phase of glycolysis; second step.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ515554; CAD56495.1; -.
DR   HSSP; P36204; 1VPE.
DR   HAMAP; MF_00145; -; 1.
DR   InterPro; IPR001576; PGK.
DR   Pfam; PF00162; PGK; 1.
DR   PIRSF; PIRSF000724; Pgk; 1.
DR   PRINTS; PR00477; PHGLYCKINASE.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Glycolysis; Kinase; Transferase.
SQ   SEQUENCE   403 AA;  42763 MW;  97CEB13D2F41BE2B CRC64;

  Query Match             15.7%;  Score 71;  DB 1;  Length 403;
  Best Local Similarity   39.5%;  Pred. No. 16;
  Matches   17;  Conservative    7;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         19 GDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDD 61
              |||  : : | || :| ||  ::|  ||||:   |   | :|
Db        110 GDVVVLENTRFQDIDNDFGKRESKNDPKLGEYWASLGDVFVND 152


RESULT 13
PGK_LACJO
ID   PGK_LACJO      STANDARD;      PRT;   403 AA.
AC   P62413;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Phosphoglycerate kinase (EC 2.7.2.3).
GN   Name=pgk; OrderedLocusNames=LJ0873;
OS   Lactobacillus johnsonii.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=33959;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 533;
RX   PubMed=14966310; DOI=10.1073/pnas.0307327101;
RA   Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA   Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA   Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT   "The genome sequence of the probiotic intestinal bacterium
RT   Lactobacillus johnsonii NCC 533.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-phospho-
CC       D-glyceroyl phosphate.
CC   -!- PATHWAY: Second phase of glycolysis; second step.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE017202; AAS08694.1; -.
DR   HAMAP; MF_00145; -; 1.
DR   PIRSF; PIRSF000724; Pgk; 1.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Complete proteome; Glycolysis; Kinase; Transferase.
SQ   SEQUENCE   403 AA;  43018 MW;  1B2F3C5671BBAB71 CRC64;

  Query Match             15.7%;  Score 71;  DB 1;  Length 403;
  Best Local Similarity   39.5%;  Pred. No. 16;
  Matches   17;  Conservative    7;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         19 GDVGQIFSCRLQDTNNFFGAGQNKRPPKLGQIGRSKRVVIEDD 61
```

```
              |||  : : | || :| ||  ::|  ||||:   |   |  :|
Db        110 GDVVVLENTRFQDIDNDFGKRESKNDPKLGEYWASLGDVFVND 152


RESULT 14
Q8QQC0
ID   Q8QQC0      PRELIMINARY;      PRT;   620 AA.
AC   Q8QQC0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   E1.
OS   Equus caballus papillomavirus type 1.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus; Equine papillomaviruses.
OX   NCBI_TaxID=191825;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ghim S.-J., Delius H., Rector A., Sundberg J.P., Jenson A.B.,
RA   Van Ranst M.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF498323; AAM19221.1; -.
DR   HSSP; P03116; 1F08.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004003; F:ATP-dependent DNA helicase activity; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   InterPro; IPR001177; Papillom_E1.
DR   Pfam; PF00519; PPV_E1_C; 1.
DR   Pfam; PF00524; PPV_E1_N; 1.
SQ   SEQUENCE   620 AA;  70618 MW;  03779145406DC888 CRC64;

  Query Match             15.7%;  Score 71;  DB 2;  Length 620;
  Best Local Similarity   37.5%;  Pred. No. 27;
  Matches   15;  Conservative    4;  Mismatches   19;  Indels    2;  Gaps    1;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQ 40
              :|  |   |::    || ||    | |  : ||  || ||
Db        582 LWRALDLSDQE--DEGDDGDPAPAFRCCTRKTNESFGGGQ 619


RESULT 15
Q8BE74
ID   Q8BE74      PRELIMINARY;      PRT;   620 AA.
AC   Q8BE74;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative replication protein E1.
OS   Equinus papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus; Equine papillomaviruses.
OX   NCBI_TaxID=185207;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22199168; PubMed=12208979;
RX   DOI=10.1128/JVI.76.19.10020-10023.2002;
RA   Terai M., DeSalle R., Burk R.D.;
RT   "Lack of canonical E6 and E7 open reading frames in bird
RT   papillomaviruses: Fringilla coelebs papillomavirus and Psittacus
RT   erithacus timneh papillomavirus.";
RL   J. Virol. 76:10020-10023(2002).
DR   EMBL; AF394740; AAN09896.1; -.
DR   HSSP; P03116; 1F08.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004003; F:ATP-dependent DNA helicase activity; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   InterPro; IPR001177; Papillom_E1.
DR   Pfam; PF00519; PPV_E1_C; 1.
DR   Pfam; PF00524; PPV_E1_N; 1.
SQ   SEQUENCE   620 AA;  70618 MW;  03779145406DC888 CRC64;

  Query Match             15.7%;  Score 71;  DB 2;  Length 620;
  Best Local Similarity   37.5%;  Pred. No. 27;
  Matches   15;  Conservative    4;  Mismatches   19;  Indels    2;  Gaps    1;

Qy          1 MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQ 40
              :|  |   |::    || ||    | |  : ||  || ||
Db        582 LWRALDLSDQE--DEGDDGDPAPAFRCCTRKTNESFGGGQ 619


Search completed: March 28, 2005, 13:01:31
Job time : 137 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 27, 2005, 21:00:22 ; Search time 13314.6 Seconds
(without alignments)
11605.570 Million cell updates/sec

Title: US-10-071-645-3
Perfect score: 3189
Sequence: 1 ccctcccctcccccctccgcc.....gaaatgccaaaaggctttta 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

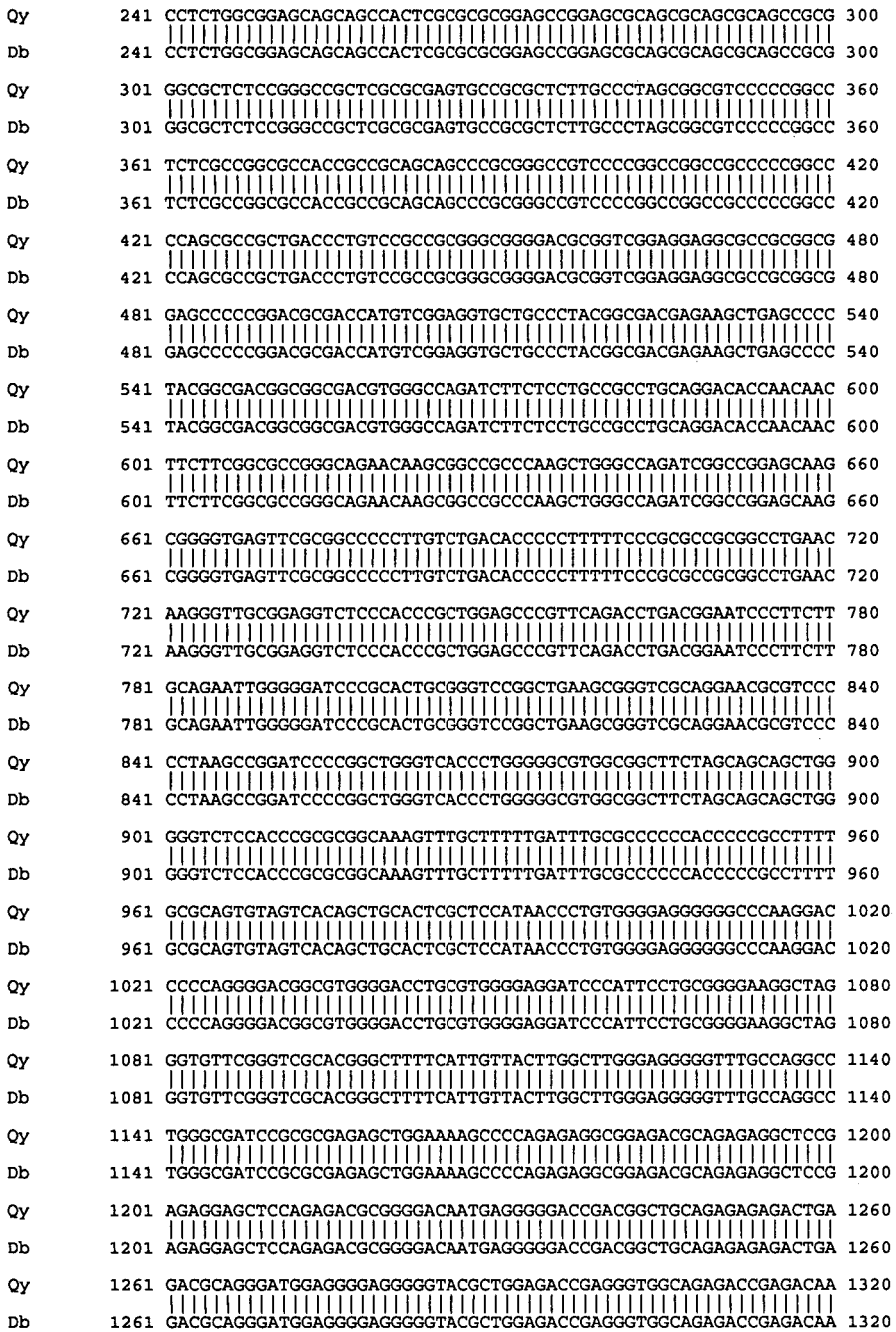
Result No.		Score	% Query Match	Length	DB	ID	Description
1		3189	100.0	3189	6	AX451461	AX451461 Sequence
c 2		3169	99.4	195076	9	AL391357	AL391357 Human DNA
3		2480	77.8	126141	2	AL356300	AL356300 Homo sapi
4		1648.4	51.7	163801	2	AC025224	AC025224 Homo sapi
c 5		1000.8	31.4	1321	6	CQ776764	CQ776764 Sequence
6		998.4	31.3	1581	6	BD276315	BD276315 MOLECULES
7		998.4	31.3	1581	6	AX049462	AX049462 Sequence
8		987.8	31.0	1377	9	BC020630	BC020630 Homo sapi
9		963.8	30.2	1876	6	AR454597	AR454597 Sequence
10		805.2	25.2	248479	10	AL807249	AL807249 Mouse DNA
c 11		734.4	23.0	170878	2	AC107845	AC107845 Mus muscu
12		713.4	22.4	259224	2	AC109006	AC109006 Rattus no
13		607.2	19.0	1104	6	AX451459	AX451459 Sequence
14		564.2	17.7	860	9	AY204901	AY204901 Homo sapi
c 15		538.8	16.9	616	9	HSA324422	AJ324422 Homo sapi
16		507	15.9	524	6	AX894800	AX894800 Sequence
17		507	15.9	524	6	BD030333	BD030333 Sequence
c 18		413.4	13.0	417	6	AR272746	AR272746 Sequence
c 19		413.4	13.0	417	6	AR276327	AR276327 Sequence
c 20		413.4	13.0	417	6	AR406602	AR406602 Sequence
c 21		413.4	13.0	417	6	AR440452	AR440452 Sequence
c 22		413.4	13.0	417	6	AR472610	AR472610 Sequence
c 23		413.4	13.0	417	6	AR543263	AR543263 Sequence
c 24		413.4	13.0	417	6	AX062848	AX062848 Sequence
c 25		413.4	13.0	417	6	AX367765	AX367765 Sequence
26		399.2	12.5	417	6	BD139338	BD139338 Extended
27		359.2	11.3	411	6	AX330168	AX330168 Sequence
28		359.2	11.3	411	6	AX332831	AX332831 Sequence
29		359.2	11.3	411	6	AX334667	AX334667 Sequence
30		359.2	11.3	411	6	AX335116	AX335116 Sequence
31		342.2	10.7	366	6	CQ704178	CQ704178 Sequence
32		319	10.0	336	6	AR563282	AR563282 Sequence
33		319	10.0	336	6	AX322029	AX322029 Sequence
34		309.2	9.7	892	11	G29976	G29976 human STS S
35		295	9.3	360	6	CQ699651	CQ699651 Sequence
36		291.2	9.1	1592	10	AY523601	AY523601 Mus muscu
37		263.2	8.3	308	6	CQ427555	CQ427555 Sequence
38		263	8.2	333	6	CQ418653	CQ418653 Sequence
39		199	6.2	1045	6	BD276322	BD276322 MOLECULES
40		199	6.2	1045	6	AX049470	AX049470 Sequence
41		177.8	5.6	672	9	HSA336572	AJ336572 Homo sapi
42		174.6	5.5	637	9	HSA336591	AJ336591 Homo sapi
43		174.6	5.5	709	9	HSA327220	AJ327220 Homo sapi
44		171.6	5.4	409	10	AF271156	AF271156 Rattus no
45		167.4	5.2	237	6	CQ727202	CQ727202 Sequence

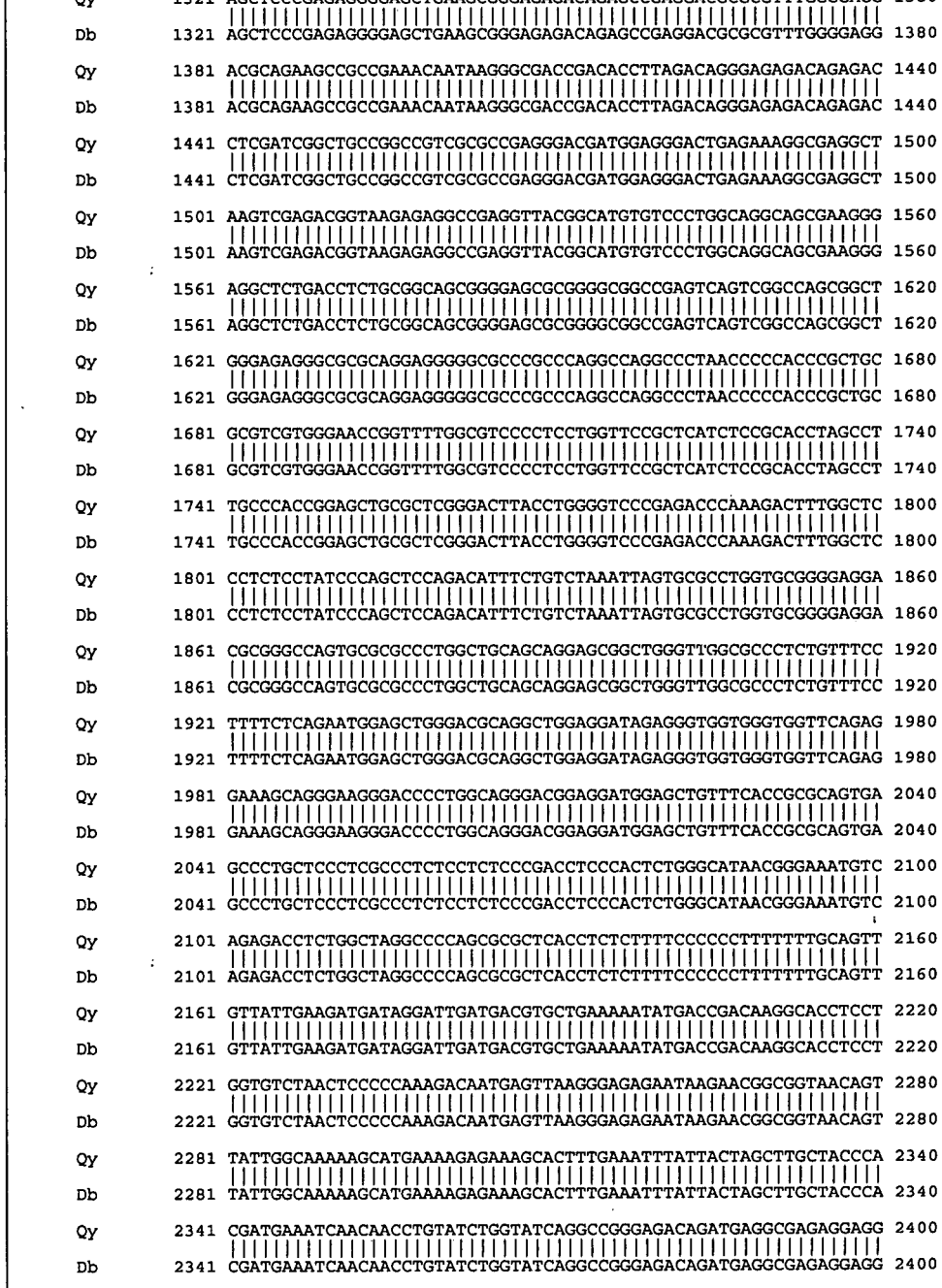
ALIGNMENTS

RESULT 1
AX451461
LOCUS AX451461 3189 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 3 from Patent WO0212279.
ACCESSION AX451461
VERSION AX451461.1 GI:21698462
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bihain,B., Bour,B. and Bougueleret,L.
TITLE Schizophrenia related gene and protein
JOURNAL Patent: WO 0212279-A 3 14-FEB-2002;
GENSET (FR)
FEATURES Location/Qualifiers
source 1. .3189
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

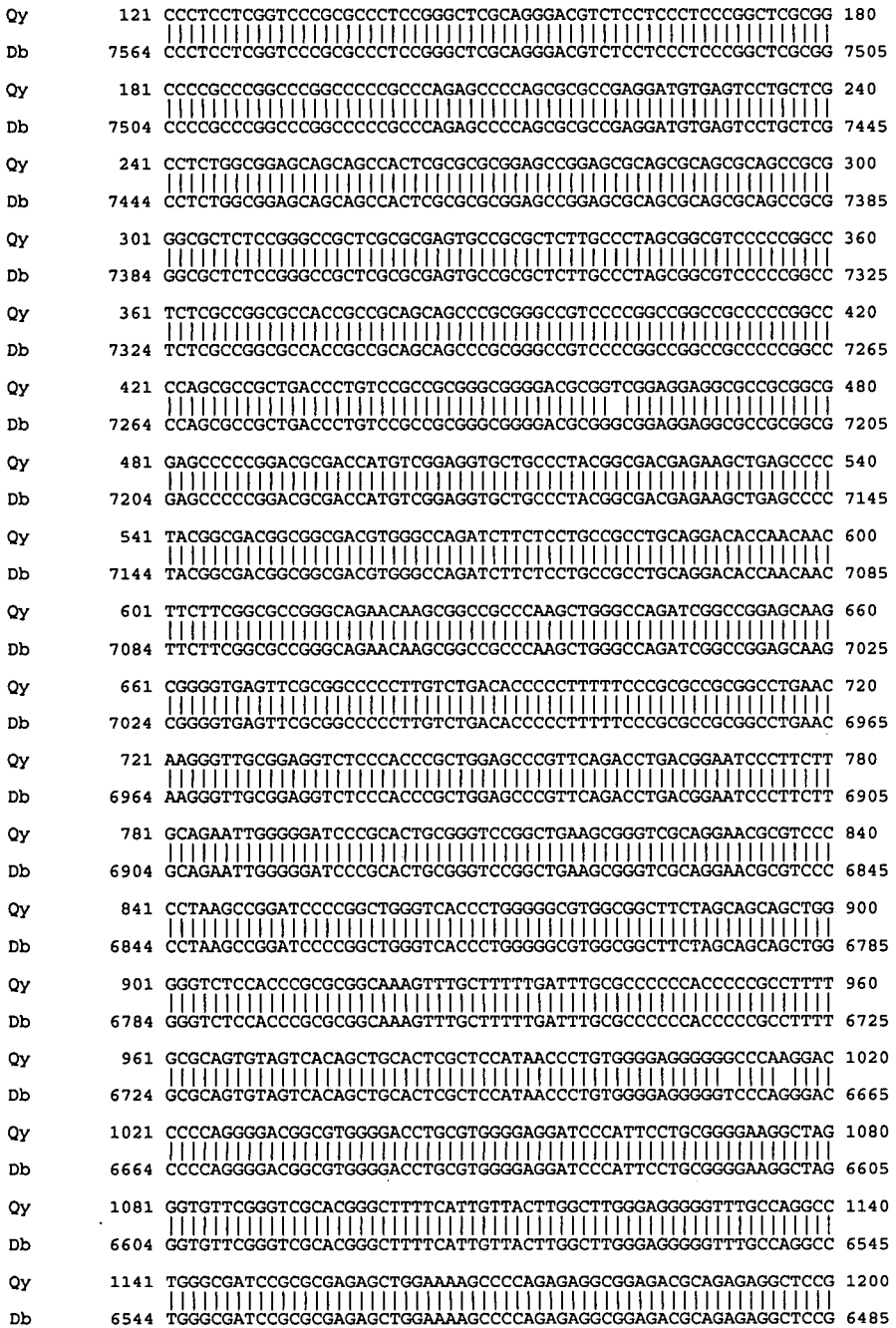
Query Match 100.0%; Score 3189; DB 6; Length 3189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

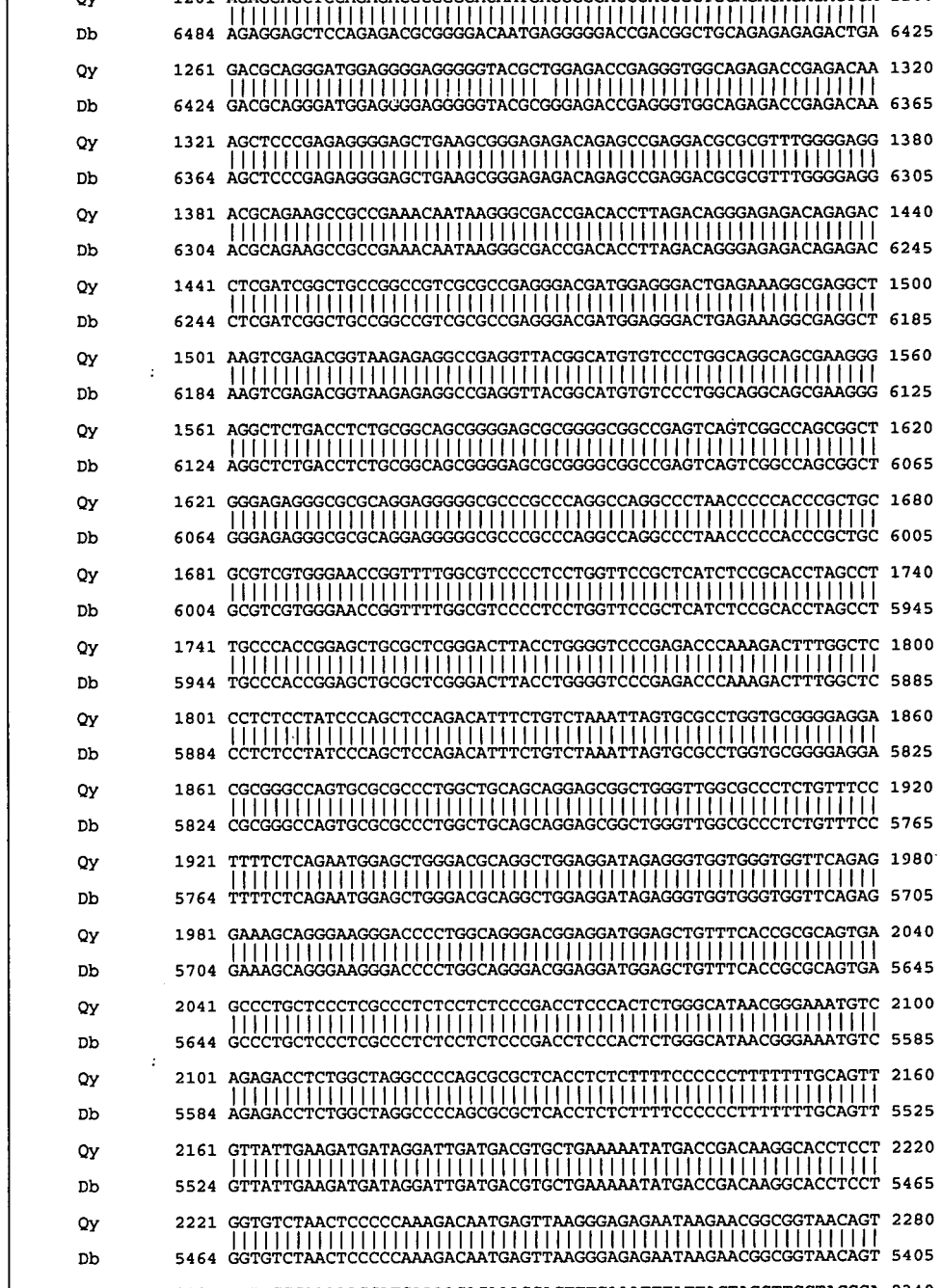
```
Qy          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120

Qy        121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180

Qy        181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
```


Qy 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Db 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Qy 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
Db 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
Qy 361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
Db 361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
Qy 421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
Db 421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
Qy 481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
Db 481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
Qy 541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
Db 541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
Qy 601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
Db 601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
Qy 661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
Db 661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
Qy 721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
Db 721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
Qy 781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
Db 781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
Qy 841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
Db 841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
Qy 901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
Db 901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
Qy 961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
Db 961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
Qy 1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
Db 1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
Qy 1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
Db 1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
Qy 1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
Db 1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
Qy 1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
Db 1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
Qy 1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320
Db 1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320


Qy 1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
Db 1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
Qy 1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
Db 1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
Qy 1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
Db 1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
Qy 1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
Db 1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
Qy 1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
Db 1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
Qy 1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
Db 1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
Qy 1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
Db 1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
Qy 1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
Db 1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
Qy 1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
Db 1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
Qy 1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
Db 1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
Qy 1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
Db 1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
Qy 1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
Db 1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
Qy 2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
Db 2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
Qy 2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
Db 2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
Qy 2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAAATATGACCGACAAGGCACCTCCT 2220
Db 2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAAATATGACCGACAAGGCACCTCCT 2220
Qy 2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280
Db 2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280
Qy 2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340
Db 2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340
Qy 2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400
Db 2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400
Qy 2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460

Qy       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520

Qy       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580

Qy       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640

Qy       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700

Qy       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760

Qy       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820

Qy       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880

Qy       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940

Qy       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000

Qy       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060

Qy       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120

Qy       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180

Qy       3181 AGGCTTTTA 3189
                 |||||||||
Db       3181 AGGCTTTTA 3189


RESULT 2
AL391357/c
LOCUS       AL391357              195076 bp    DNA     linear   PRI 25-OCT-2001
DEFINITION  Human DNA sequence from clone RP11-401M16 on chromosome 1, complete
            sequence.
ACCESSION   AL391357
VERSION     AL391357.20  GI:16501137
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 195076)
  AUTHORS   Kimberley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 26, 2001 this sequence version replaced gi:16304938.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-401M16 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-401M16 The true
            left end of clone RP11-72P17 is at 108464 in this sequence. The
            true right end of clone RP4-749H3 is at 12964 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .195076
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-401M16"
                     /clone_lib="RPCI-11.2"
     misc_feature    5777. .8825
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    29731. .30197
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    74209. .75906
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    75757. .75801
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    145474. .145654
                     /note="Sequence from overlapping clone RP11-72P17
                     (AL365439). Assembly confirmed by restriction digest."
     misc_feature    155179. .155865
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    182831. .183552
                     /note="CpG island"
                     /evidence=not_experimental
ORIGIN

  Query Match             99.4%;  Score 3169;  DB 9;  Length 195076;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3183;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7684 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 7625

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7624 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 7565
```


Qy 121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
Db 7564 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 7505
Qy 181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
Db 7504 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 7445
Qy 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Db 7444 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 7385
Qy 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
Db 7384 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 7325
Qy 361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
Db 7324 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 7265
Qy 421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
Db 7264 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGGCGGAGGAGGCGCCGCGGCG 7205
Qy 481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
Db 7204 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 7145
Qy 541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
Db 7144 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 7085
Qy 601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
Db 7084 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 7025
Qy 661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
Db 7024 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 6965
Qy 721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
Db 6964 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 6905
Qy 781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
Db 6904 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 6845
Qy 841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
Db 6844 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 6785
Qy 901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
Db 6784 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 6725
Qy 961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
Db 6724 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGTCCCAGGGAC 6665
Qy 1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
Db 6664 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 6605
Qy 1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
Db 6604 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 6545
Qy 1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
Db 6544 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 6485


Qy 1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
Db 6484 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 6425
Qy 1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320
Db 6424 GACGCAGGGATGGAGGGGAGGGGGTACGCGGGAGACCGAGGGTGGCAGAGACCGAGACAA 6365
Qy 1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
Db 6364 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 6305
Qy 1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
Db 6304 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 6245
Qy 1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
Db 6244 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 6185
Qy 1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
Db 6184 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 6125
Qy 1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
Db 6124 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 6065
Qy 1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
Db 6064 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 6005
Qy 1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
Db 6004 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 5945
Qy 1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
Db 5944 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 5885
Qy 1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
Db 5884 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 5825
Qy 1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
Db 5824 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 5765
Qy 1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
Db 5764 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 5705
Qy 1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
Db 5704 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 5645
Qy 2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
Db 5644 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 5585
Qy 2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
Db 5584 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 5525
Qy 2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAAATATGACCGACAAGGCACCTCCT 2220
Db 5524 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAAATATGACCGACAAGGCACCTCCT 5465
Qy 2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280
Db 5464 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 5405
Qy 2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340

```
Db       5404 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 5345

Qy       2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5344 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 5285

Qy       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5284 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 5225

Qy       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5224 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 5165

Qy       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5164 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 5105

Qy       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5104 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 5045

Qy       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5044 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 4985

Qy       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4984 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 4925

Qy       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4924 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 4865

Qy       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4864 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 4805

Qy       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4804 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 4745

Qy       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4744 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 4685

Qy       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4684 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 4625

Qy       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4624 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 4565

Qy       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAATGCCAA 3179
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       4564 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAATGCCAA 4505

Qy       3180 AAGGCTTTT 3188
             ||||| |||
Db       4504 AAGGCATTT 4496


RESULT 3
AL356300
LOCUS       AL356300              126141 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP4-658J7 map p36.11-36.13, 12
            unordered pieces.
ACCESSION   AL356300
VERSION     AL356300.5  GI:11137765
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plumb,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 10, 2000 this sequence version replaced gi:9797484.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ658J7
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 120052 bases at least Q40
            Consensus quality: 122543 bases at least Q30
            Consensus quality: 123707 bases at least Q20
            Insert size: 125041; sum-of-contigs
            Insert size: 147762; 1.2% error; agarose-fp
            Quality coverage: 3.57x in Q20 bases; sum-of-contigs Quality
            coverage: 3.10x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    16101: contig of 16101 bp in length
            *    16102    16201: gap of 100 bp
            *    16202    19430: contig of 3229 bp in length
            *    19431    19530: gap of 100 bp
            *    19531    46459: contig of 26929 bp in length
            *    46460    46559: gap of 100 bp
            *    46560    51855: contig of 5296 bp in length
            *    51856    51955: gap of 100 bp
            *    51956    58425: contig of 6470 bp in length
            *    58426    58525: gap of 100 bp
            *    58526    69865: contig of 11340 bp in length
            *    69866    69965: gap of 100 bp
            *    69966    73734: contig of 3769 bp in length
            *    73735    73834: gap of 100 bp
            *    73835    91543: contig of 17709 bp in length
            *    91544    91643: gap of 100 bp
            *    91644    93828: contig of 2185 bp in length
            *    93829    93928: gap of 100 bp
            *    93929   102232: contig of 8304 bp in length
            *   102233   102332: gap of 100 bp
            *   102333   109944: contig of 7612 bp in length
            *   109945   110044: gap of 100 bp
            *   110045   126141: contig of 16097 bp in length.
FEATURES             Location/Qualifiers
     source          1. .126141
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p36.11-36.13"
                     /clone="RP4-658J7"
                     /clone_lib="RPCI-4"
     misc_feature    1. .16101
                     /note="assembly_fragment:00015
                     fragment_chain:1"
     misc_feature    16202. .19430
```

```
                     /note="assembly_fragment:00124
                     fragment_chain:1"
     misc_feature    19531. .46459
                     /note="assembly_fragment:00873
                     fragment_chain:1"
     misc_feature    46560. .51855
                     /note="assembly_fragment:00305
                     fragment_chain:2"
     misc_feature    51956. .58425
                     /note="assembly_fragment:00664
                     fragment_chain:2"
     misc_feature    58526. .69865
                     /note="assembly_fragment:00698
                     fragment_chain:3"
     misc_feature    69966. .73734
                     /note="assembly_fragment:01008
                     fragment_chain:3"
     misc_feature    73835. .91543
                     /note="assembly_fragment:00799
                     fragment_chain:4"
     misc_feature    91644. .93828
                     /note="assembly_fragment:00579
                     fragment_chain:4"
     misc_feature    93929. .102232
                     /note="assembly_fragment:00217"
     misc_feature    102333. .109944
                     /note="assembly_fragment:00284"
     misc_feature    110045. .126141
                     /note="assembly_fragment:01133"
ORIGIN

  Query Match             77.8%;  Score 2480;  DB 2;  Length 126141;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2494;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        690 CACCCCCTTTTTCCCGCGCCGCGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGG 749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73835 CACCCCCTTTTTCCCGCGCCGCGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGG 73894

Qy        750 AGCCCGTTCAGACCTGACGGAATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGT 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73895 AGCCCGTTCAGACCTGACGGAATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGT 73954

Qy        810 CCGGCTGAAGCGGGTCGCAGGAACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73955 CCGGCTGAAGCGGGTCGCAGGAACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCC 74014

Qy        870 TGGGGGCGTGGCGGCTTCTAGCAGCAGCTGGGGGTCTCCACCCGCGCGGCAAAGTTTGCT 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74015 TGGGGGCGTGGCGGCTTCTAGCAGCAGCTGGGGGTCTCCACCCGCGCGGCAAAGTTTGCT 74074

Qy        930 TTTTGATTTGCGCCCCCCACCCCCGCCTTTTGCGCAGTGTAGTCACAGCTGCACTCGCTC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74075 TTTTGATTTGCGCCCCCCACCCCCGCCTTTTGCGCAGTGTAGTCACAGCTGCACTCGCTC 74134

Qy        990 CATAACCCTGTGGGGAGGGGGGCCCAAGGACCCCCAGGGGACGGCGTGGGGACCTGCGTG 1049
              |||||||||| |||||||||| |||| ||||||||||||||||||||||||||||||||
Db      74135 CATAACCCTGGGGGGAGGGGGTCCCAGGGACCCCCAGGGGACGGCGTGGGGACCTGCGTG 74194

Qy       1050 GGGAGGATCCCATTCCTGCGGGGAAGGCTAGGGTGTTCGGGTCGCACGGGCTTTTCATTG 1109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74195 GGGAGGATCCCATTCCTGCGGGGAAGGCTAGGGTGTTCGGGTCGCACGGGCTTTTCATTG 74254

Qy       1110 TTACTTGGCTTGGGAGGGGGTTTGCCAGGCCTGGGCGATCCGCGCGAGAGCTGGAAAAGC 1169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74255 TTACTTGGCTTGGGAGGGGGTTTGCCAGGCCTGGGCGATCCGCGCGAGAGCTGGAAAAGC 74314

Qy       1170 CCCAGAGAGGCGGAGACGCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGGGGACAATG 1229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74315 CCCAGAGAGGCGGAGACGCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGGGGACAATG 74374

Qy       1230 AGGGGGACCGACGGCTGCAGAGAGAGACTGAGACGCAGGGATGGAGGGGAGGGGGTACGC 1289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      74375 AGGGGGACCGACGGCTGCAGAGAGAGACTGAGACGCAGGGATGGAGGGGAGGGGGTACGC 74434

Qy       1290 TGGAGACCGAGGGTGGCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAG 1349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74435 GGGAGACCGAGGGTGGCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAG 74494

Qy       1350 AGACAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAATAAGGGCG 1409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74495 AGACAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAATAAGGGCG 74554

Qy       1410 ACCGACACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCCGTCGCGCCGAG 1469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74555 ACCGACACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCCGTCGCGCCGAG 74614

Qy       1470 GGACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTT 1529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74615 GGACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTT 74674

Qy       1530 ACGGCATGTGTCCCTGGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCAGCGGGGAGC 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74675 ACGGCATGTGTCCCTGGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCAGCGGGGAGC 74734

Qy       1590 GCGGGGCGGCCGAGTCAGTCGGCCAGCGGCTGGGAGAGGGCGCGCAGGAGGGGGCGCCCG 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74735 GCGGGGCGGCCGAGTCAGTCGGCCAGCGGCTGGGAGAGGGCGCGCAGGAGGGGGCGCCCG 74794

Qy       1650 CCCAGGCCAGGCCCTAACCCCCACCCGCTGCGCGTCGTGGGAACCGGTTTTGGCGTCCCC 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74795 CCCAGGCCAGGCCCTAACCCCCACCCGCTGCGCGTCGTGGGAACCGGTTTTGGCGTCCCC 74854

Qy       1710 TCCTGGTTCCGCTCATCTCCGCACCTAGCCTTGCCCACCGGAGCTGCGCTCGGGACTTAC 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74855 TCCTGGTTCCGCTCATCTCCGCACCTAGCCTTGCCCACCGGAGCTGCGCTCGGGACTTAC 74914

Qy       1770 CTGGGGTCCCGAGACCCAAAGACTTTGGCTCCCTCTCCTATCCCAGCTCCAGACATTTCT 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74915 CTGGGGTCCCGAGACCCAAAGACTTTGGCTCCCTCTCCTATCCCAGCTCCAGACATTTCT 74974

Qy       1830 GTCTAAATTAGTGCGCCTGGTGCGGGGAGGACGCGGGCCAGTGCGCGCCCTGGCTGCAGC 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74975 GTCTAAATTAGTGCGCCTGGTGCGGGGAGGACGCGGGCCAGTGCGCGCCCTGGCTGCAGC 75034

Qy       1890 AGGAGCGGCTGGGTTGGCGCCCTCTGTTTCCTTTTCTCAGAATGGAGCTGGGACGCAGGC 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75035 AGGAGCGGCTGGGTTGGCGCCCTCTGTTTCCTTTTCTCAGAATGGAGCTGGGACGCAGGC 75094

Qy       1950 TGGAGGATAGAGGGTGGTGGGTGGTTCAGAGGAAAGCAGGGAAGGGACCCCTGGCAGGGA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75095 TGGAGGATAGAGGGTGGTGGGTGGTTCAGAGGAAAGCAGGGAAGGGACCCCTGGCAGGGA 75154

Qy       2010 CGGAGGATGGAGCTGTTTCACCGCGCAGTGAGCCCTGCTCCCTCGCCCTCTCCTCTCCCG 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75155 CGGAGGATGGAGCTGTTTCACCGCGCAGTGAGCCCTGCTCCCTCGCCCTCTCCTCTCCCG 75214

Qy       2070 ACCTCCCACTCTGGGCATAACGGGAAATGTCAGAGACCTCTGGCTAGGCCCCAGCGCGCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75215 ACCTCCCACTCTGGGCATAACGGGAAATGTCAGAGACCTCTGGCTAGGCCCCAGCGCGCT 75274

Qy       2130 CACCTCTCTTTTCCCCCCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGATGACGT 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75275 CACCTCTCTTTTCCCCCCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGATGACGT 75334

Qy       2190 GCTGAAAAATATGACCGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75335 GCTGAAAAATATGACCGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGT 75394

Qy       2250 TAAGGGAGAGAATAAGAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGC 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75395 TAAGGGAGAGAATAAGAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGC 75454

Qy       2310 ACTTTGAAATTTATTACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTAT 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75455 ACTTTGAAATTTATTACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTAT 75514
```

```
Qy       2370 CAGGCCGGGAGACAGATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCC 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75515 CAGGCCGGGAGACAGATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCC 75574

Qy       2430 TCTGCAAAAATAAAAATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATA 2489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75575 TCTGCAAAAATAAAAATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATA 75634

Qy       2490 TACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCC 2549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75635 TACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCC 75694

Qy       2550 ATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAA 2609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75695 ATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAA 75754

Qy       2610 TGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCC 2669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75755 TGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCC 75814

Qy       2670 AATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGT 2729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75815 AATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGT 75874

Qy       2730 CCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCT 2789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75875 CCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCT 75934

Qy       2790 CTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAAC 2849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75935 CTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAAC 75994

Qy       2850 AAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGT 2909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75995 AAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGT 76054

Qy       2910 AAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACAT 2969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76055 AAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACAT 76114

Qy       2970 TTTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACACGAAAAACA 3029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76115 TTTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACGAAAAACA 76174

Qy       3030 AAGAAAAAAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCT 3089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76175 AAGAAAAAAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCT 76234

Qy       3090 GTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAAT 3149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76235 GTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAAT 76294

Qy       3150 CACAAGAGTGCCAAAT-TCTTGAAATGCCAAAAGGCTTTT 3188
             ||||||||||||||| |||||||||||||||||| |||
Db      76295 CACAAGAGTGCCAAATATCTTGAAATGCCAAAAGGCATTT 76334
```

```
RESULT 4
AC025224
LOCUS       AC025224               163801 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-4L9, WORKING DRAFT SEQUENCE,
            31 unordered pieces.
ACCESSION   AC025224
VERSION     AC025224.3  GI:7596970
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163801)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163801)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Apr 19, 2000 this sequence version replaced gi:7321912.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0004L09
            -------------- Summary Statistics --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 145820 bases at least Q40
            Consensus quality: 150999 bases at least Q30
            Consensus quality: 153311 bases at least Q20
            Insert size: 185000; agarose-fp
            Insert size: 160801; sum-of-contigs
            Quality coverage: 3.57 in Q20 bases; agarose-fp
            Quality coverage: 4.25 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 31 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1640: contig of 1640 bp in length
            *     1641     1740: gap of unknown length
            *     1741     3443: contig of 1703 bp in length
            *     3444     3543: gap of unknown length
            *     3544     4790: contig of 1247 bp in length
            *     4791     4890: gap of unknown length
            *     4891     6777: contig of 1887 bp in length
            *     6778     6877: gap of unknown length
            *     6878     8114: contig of 1237 bp in length
            *     8115     8214: gap of unknown length
            *     8215     9912: contig of 1698 bp in length
            *     9913    10012: gap of unknown length
            *    10013    11754: contig of 1742 bp in length
            *    11755    11854: gap of unknown length
            *    11855    14058: contig of 2204 bp in length
            *    14059    14158: gap of unknown length
            *    14159    18450: contig of 4292 bp in length
            *    18451    18550: gap of unknown length
            *    18551    22470: contig of 3920 bp in length
            *    22471    22570: gap of unknown length
            *    22571    26337: contig of 3767 bp in length
            *    26338    26437: gap of unknown length
            *    26438    30118: contig of 3681 bp in length
            *    30119    30218: gap of unknown length
            *    30219    33161: contig of 2943 bp in length
            *    33162    33261: gap of unknown length
            *    33262    37679: contig of 4418 bp in length
            *    37680    37779: gap of unknown length
            *    37780    41155: contig of 3376 bp in length
            *    41156    41255: gap of unknown length
            *    41256    45149: contig of 3894 bp in length
            *    45150    45249: gap of unknown length
            *    45250    50155: contig of 4906 bp in length
            *    50156    50255: gap of unknown length
            *    50256    53871: contig of 3616 bp in length
            *    53872    53971: gap of unknown length
```

```
            *    53972    58814: contig of 4843 bp in length
            *    58815    58914: gap of unknown length
            *    58915    63420: contig of 4506 bp in length
            *    63421    63520: gap of unknown length
            *    63521    67773: contig of 4253 bp in length
            *    67774    67873: gap of unknown length
            *    67874    73896: contig of 6023 bp in length
            *    73897    73996: gap of unknown length
            *    73997    79302: contig of 5306 bp in length
            *    79303    79402: gap of unknown length
            *    79403    85385: contig of 5983 bp in length
            *    85386    85485: gap of unknown length
            *    85486    93969: contig of 8484 bp in length
            *    93970    94069: gap of unknown length
            *    94070   102628: contig of 8559 bp in length
            *   102629   102728: gap of unknown length
            *   102729   113357: contig of 10629 bp in length
            *   113358   113457: gap of unknown length
            *   113458   123217: contig of 9760 bp in length
            *   123218   123317: gap of unknown length
            *   123318   132911: contig of 9594 bp in length
            *   132912   133011: gap of unknown length
            *   133012   146300: contig of 13289 bp in length
            *   146301   146400: gap of unknown length
            *   146401   163801: contig of 17401 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163801
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-4L9"
     misc_feature    1. .1640
                     /note="assembly_name:Contig12"
     misc_feature    1741. .3443
                     /note="assembly_name:Contig13"
     misc_feature    3544. .4790
                     /note="assembly_name:Contig14"
     misc_feature    4891. .6777
                     /note="assembly_name:Contig15"
     misc_feature    6878. .8114
                     /note="assembly_name:Contig16"
     misc_feature    8215. .9912
                     /note="assembly_name:Contig17"
     misc_feature    10013. .11754
                     /note="assembly_name:Contig18"
     misc_feature    11855. .14058
                     /note="assembly_name:Contig19
                     clone_end:SP6
                     vector_side:left"
     misc_feature    14159. .18450
                     /note="assembly_name:Contig20"
     misc_feature    18551. .22470
                     /note="assembly_name:Contig21"
     misc_feature    22571. .26337
                     /note="assembly_name:Contig22"
     misc_feature    26438. .30118
                     /note="assembly_name:Contig23
                     clone_end:T7
                     vector_side:right"
     misc_feature    30219. .33161
                     /note="assembly_name:Contig24"
     misc_feature    33262. .37679
                     /note="assembly_name:Contig25"
     misc_feature    37780. .41155
                     /note="assembly_name:Contig26"
     misc_feature    41256. .45149
                     /note="assembly_name:Contig27"
     misc_feature    45250. .50155
                     /note="assembly_name:Contig28"
     misc_feature    50256. .53871
                     /note="assembly_name:Contig29"
     misc_feature    53972. .58814
                     /note="assembly_name:Contig30"
     misc_feature    58915. .63420
                     /note="assembly_name:Contig31"
     misc_feature    63521. .67773
                     /note="assembly_name:Contig32"
     misc_feature    67874. .73896
                     /note="assembly_name:Contig33"
     misc_feature    73997. .79302
                     /note="assembly_name:Contig34"
     misc_feature    79403. .85385
                     /note="assembly_name:Contig35"
     misc_feature    85486. .93969
                     /note="assembly_name:Contig36"
     misc_feature    94070. .102628
                     /note="assembly_name:Contig37"
     misc_feature    102729. .113357
                     /note="assembly_name:Contig38"
     misc_feature    113458. .123217
                     /note="assembly_name:Contig39"
     misc_feature    123318. .132911
                     /note="assembly_name:Contig40"
     misc_feature    133012. .146300
                     /note="assembly_name:Contig41"
     misc_feature    146401. .163801
                     /note="assembly_name:Contig42"
ORIGIN

  Query Match           51.7%;  Score 1648.4;  DB 2;  Length 163801;
  Best Local Similarity   92.9%;  Pred. No. 4.3e-221;
  Matches 1755;  Conservative    0;  Mismatches  121;  Indels   14;  Gaps    8;

Qy       1305 GCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGA 1364
                | |
Db      85381 GTATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85440

Qy       1365 CGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGAC 1424
                                                                ||||||||||||||||
Db      85441 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACCGACACCTTAGAC 85500

Qy       1425 AGGGAGAGACAGAGACCTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGAC 1484
                |||||||||||||||||||||||||||||||||| | ||||||    ||||||||||||||||
Db      85501 AGGGAGAGACAGAGACCTCGATCGGCTGCCGGGC-TCGCGCGAGAGGACGATGGAGGGAC 85559

Qy       1485 TGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCT 1544
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85560 TGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCT 85619

Qy       1545 GGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGT 1604
                |||||||||||||||||||||||||||||||||||| ||||||||    |||| ||||||
Db      85620 GGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCA-CGGGGAGCCGTAGGCGACCGAGT 85678

Qy       1605 CAGTCGGCCAGCGGCTGGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCT 1664
                ||||| || ||||||||||||||||||||||||||||||    ||  ||||||||||||||
Db      85679 CAGTC-GCGAGCGGCTGGGAGAGGGCGCGCAGGAGGGGG---CCTTCCAGGCCAGGCCCT 85734

Qy       1665 AACCCCCACCCGCTGCGCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCA 1724
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      85735 AACCCCCACCCGCTGCGC-TCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCA 85793

Qy       1725 TCTCCGCACCTAGCCTTGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGAC 1784
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85794 TCTCCGCACCTAGCCTTGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGA- 85852

Qy       1785 CCAAAGACTTTGGCTCCCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCG 1844
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85853 CCAAAGACTTTGGCTCCCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCG 85912

Qy       1845 CCTGGTGCGGGGAGGACGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTT 1904
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85913 CCTGGTGCGGGGAGGACGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTT 85972

Qy       1905 GGCGCCCTCTGTTTCCTTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGT 1964
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      85973 GGCGCCCTCTGTTTCCTTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGT 86032

Qy       1965 GGTGGGTGGTTCAGAGGAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTG 2024
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||||
Db      86033 GGTGGGTGGTTCAGAGGAAAGCAGGGAAGGGACCCCTGGCAGGTACGGAGTATGGAGCTG 86092

Qy       2025 TTTCACCGCGCAGTGAGCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGG 2084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86093 TTTCACCGCGCAGTGAGCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGG 86152

Qy       2085 CATAACGGGAAATGTCAGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCC 2144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86153 CATAACGGGAAATGTCAGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCC 86212

Qy       2145 CCCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGAC 2204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86213 CCCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGAC 86272

Qy       2205 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 2264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86273 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 86332

Qy       2265 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 2324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86333 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 86392

Qy       2325 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 2384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86393 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 86452

Qy       2385 ATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAA 2444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86453 ATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAA 86512

Qy       2445 ATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAA 2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86513 ATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAA 86572

Qy       2505 ATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATA 2564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86573 ATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATA 86632

Qy       2565 CGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACC 2624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86633 CGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACC 86692

Qy       2625 AGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATT 2684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86693 AGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATT 86752

Qy       2685 TAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTT 2744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86753 TAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTT 86812

Qy       2745 TTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGT 2804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86813 TTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGT 86872

Qy       2805 TTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTT 2864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86873 TTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTT 86932

Qy       2865 CTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCT 2924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86933 CTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCT 86992

Qy       2925 TTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAA 2984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86993 TTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAA 87052

Qy       2985 AACACACACACACACACACACACACACACACACAC-----ACGAAAAACAAAGAAAAAAA 3039
              ||||||||||||||||||||||||||||||||||||     |||||||| ||||||||||
Db      87053 AACACACACACACACACACACACACACACACACACACACGACGAAAAACGAAGAAAAAAA 87112

Qy       3040 TGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCT 3099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      87113 TGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGATTATTTCT 87172

Qy       3100 CTAATATTATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTG 3159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87173 CTAATATTATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTG 87232

Qy       3160 CCAAAT-TCTTGAAATGCCAAAAGGCTTTT 3188
              |||||| |||||||||||||||||| |||
Db      87233 CCAAATATCTTGAAATGCCAAAAGGCATTT 87262
```

RESULT 5
CQ776764/c

```
LOCUS       CQ776764               1321 bp    DNA     linear   PAT 11-MAR-2004
DEFINITION  Sequence 450 from Patent EP1394274.
ACCESSION   CQ776764
VERSION     CQ776764.1  GI:45380154
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.
  TITLE     Methods of testing for bronchial asthma or chronic obstructive
            pulmonary disease
  JOURNAL   Patent: EP 1394274-A 450 03-MAR-2004;
            Genox Research, Inc. (JP)
FEATURES             Location/Qualifiers
     source          1. .1321
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 31.4%; Score 1000.8; DB 6; Length 1321;
Best Local Similarity 99.1%; Pred. No. 1.9e-130;
Matches 1029; Conservative 0; Mismatches 2; Indels 7; Gaps 2;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1230

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1170

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1110

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1050

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 990

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 930

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 870

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 810
```

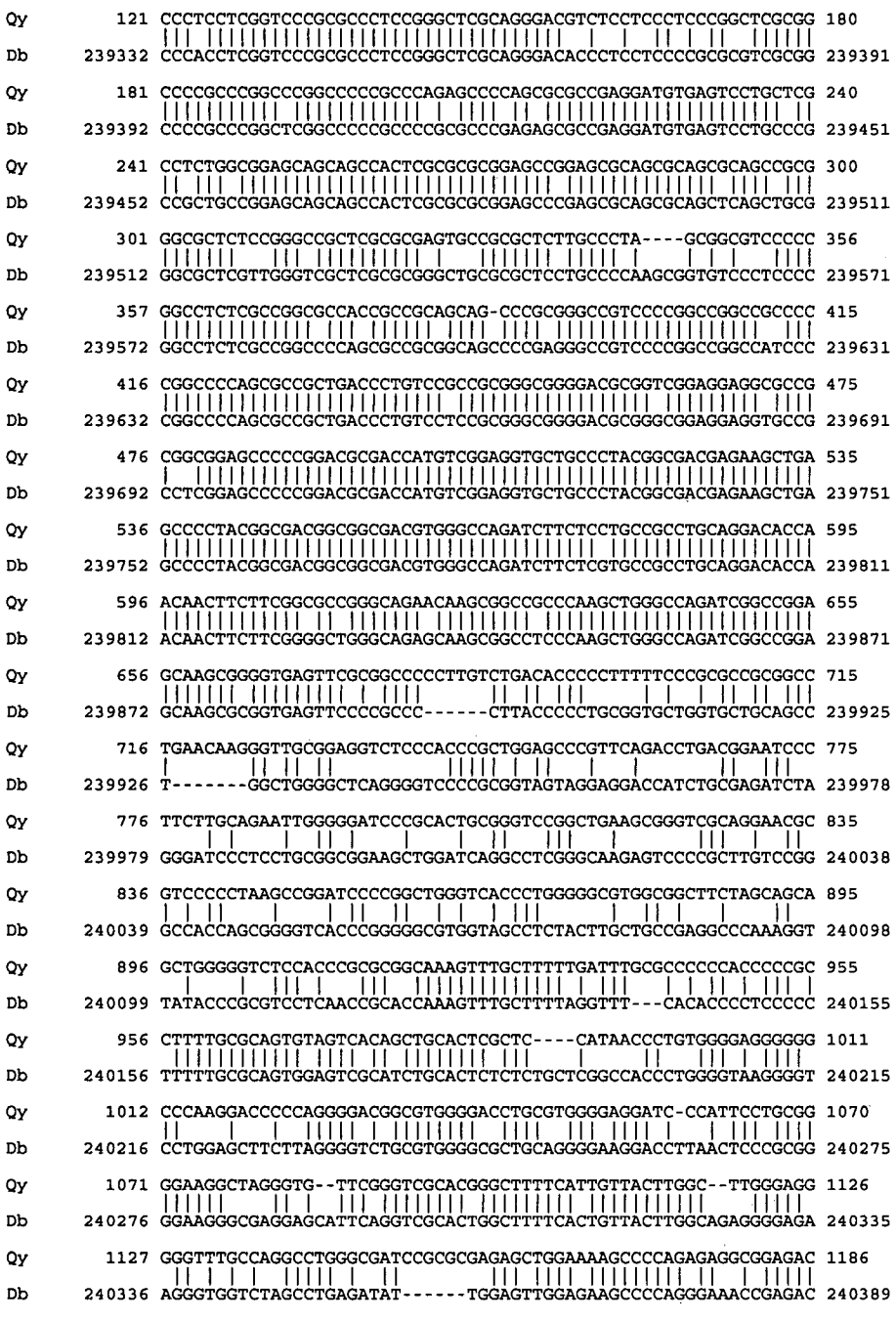
```
Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 750

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 690

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 630

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 570

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        569 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 510

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        509 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 450

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 390

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 330

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        329 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 270

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db        269 AAATGCCAAAAGGCATTT 252


RESULT 6
BD276315
LOCUS       BD276315               1581 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  MOLECULES EXPRESSED IN HIPPOCAMPUS.
ACCESSION   BD276315
VERSION     BD276315.1  GI:33086083
KEYWORDS    JP 2002543833-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1581)
  AUTHORS   Baughn,M.R., Lal,P., Yue,H., Tang,T.Y., Azimzai,Y. and Kaser,M.R.
  TITLE     MOLECULES EXPRESSED IN HIPPOCAMPUS
  JOURNAL   Patent: JP 2002543833-A 1 24-DEC-2002;
            INCYTE GENOMICS INC,Matthew R KASER,Preeti LAL,Henry YUE, Tom Y
            TANG, Mariah R BAUGHN,Yalda AZIMZAI
COMMENT     OS   Homo sapiens
            PN   JP 2002543833-A/1
            PD   24-DEC-2002
            PF   10-MAY-2000 JP 2000618442
            PR   17-MAY-1999 US    09/313300
            PI   mariah r baughn,preeti lal,henry yue,tom y tang,yalda azimzai,
            PI   matthew r kaser
            CC   This description about <220> can't be interpreted CC   <220>
            CC   <221> misc feature
            CC   <223> Incyte ID No.: 239240
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .1581
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match           31.3%;  Score 998.4;  DB 6;  Length 1581;
  Best Local Similarity   98.8%;  Pred. No. 4.2e-130;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels   11;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 263

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 323

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 383

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 443

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 503

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 563

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 623

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 683

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 743

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 803

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 863

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 923

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 983

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA----------AC 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||          ||
Db        984 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1043

Qy       2988 ACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 3047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 ACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 1103

Qy       3048 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 3107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1104 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 1163

Qy       3108 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-T 3166
```

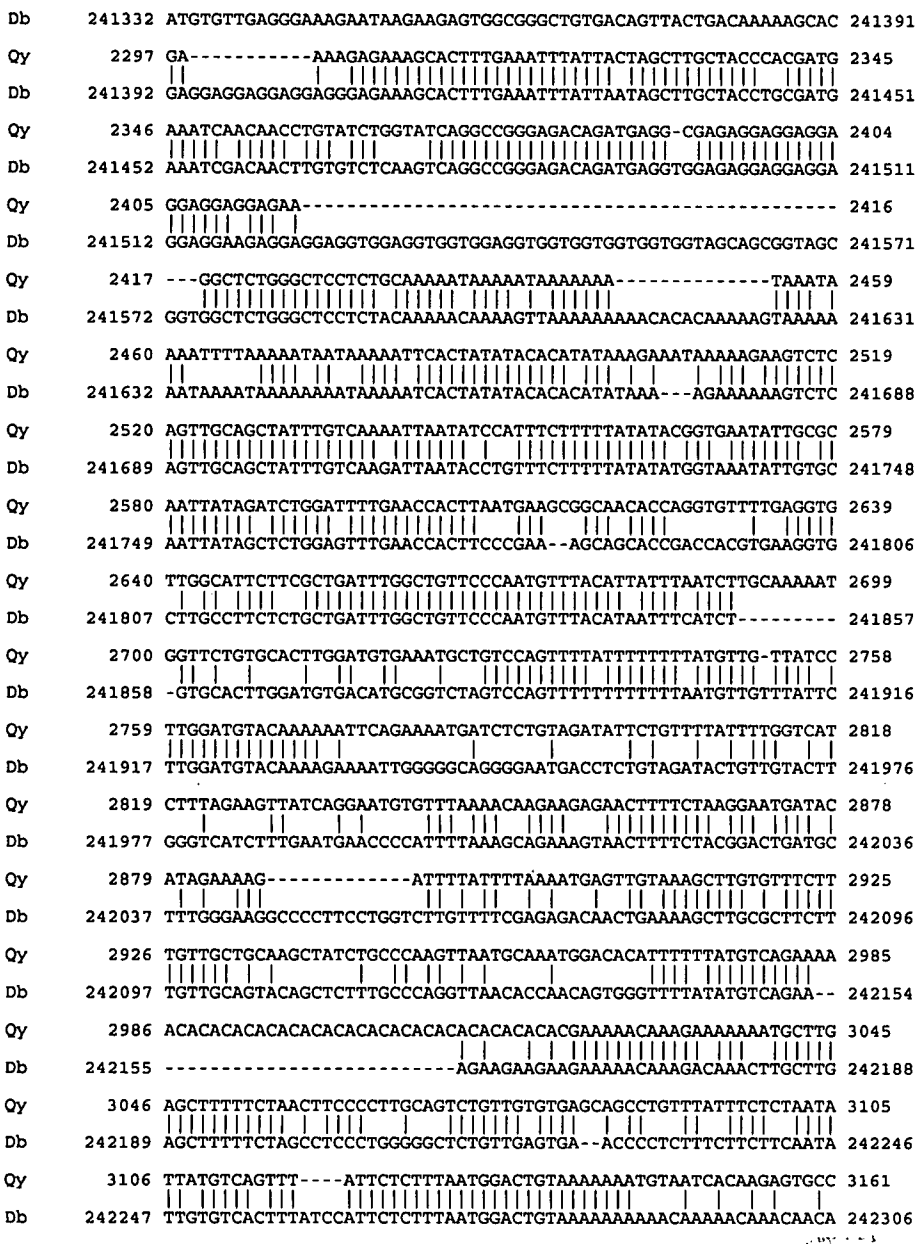
```
Db       1164 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATAT 1223

Qy       3167 CTTGAAATGCCAAAAGGCTTTT 3188
Db       1224 CTTGAAATGCCAAAAGGCATTT 1245


RESULT 7
AX049462
LOCUS       AX049462               1581 bp    DNA     linear   PAT 12-JAN-2001
DEFINITION  Sequence 1 from Patent WO0070036.
ACCESSION   AX049462
VERSION     AX049462.1  GI:12226181
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kaser,M.R., Lal,P., Yue,H., Tang,Y.T., Baughn,M.R. and Azimzai,Y.
  TITLE     Genes expressed in hippocampus
  JOURNAL   Patent: WO 0070036-A 1 23-NOV-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1581
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No.: 239240"
ORIGIN

  Query Match           31.3%;  Score 998.4;  DB 6;  Length 1581;
  Best Local Similarity   98.8%;  Pred. No. 4.2e-130;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels   11;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
Db        204 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 263

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
Db        264 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 323

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
Db        324 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 383

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
Db        384 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 443

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
Db        444 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 503

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
Db        504 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 563

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
Db        564 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 623

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
Db        624 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 683

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
Db        684 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 743

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
Db        744 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 803

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
Db        804 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 863

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
Db        864 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 923

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
Db        924 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 983

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA----------AC 2987
Db        984 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1043

Qy       2988 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 3047
Db       1044 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 1103

Qy       3048 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 3107
Db       1104 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 1163

Qy       3108 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-T 3166
Db       1164 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATAT 1223

Qy       3167 CTTGAAATGCCAAAAGGCTTTT 3188
Db       1224 CTTGAAATGCCAAAAGGCATTT 1245


RESULT 8
BC020630
LOCUS       BC020630               1377 bp    mRNA    linear   PRI 08-OCT-2003
DEFINITION  Homo sapiens calcium/calmodulin-dependent protein kinase II, mRNA
            (cDNA clone MGC:22256 IMAGE:4703846), complete cds.
ACCESSION   BC020630
VERSION     BC020630.1  GI:18089055
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1377)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1377)
```

```
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 37 Row: b Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 20591468.
FEATURES             Location/Qualifiers
     source          1. .1377
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:22256 IMAGE:4703846"
                     /tissue_type="Lung"
                     /clone_lib="NIH_MGC_77"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1. .1377
                     /gene="CaMKIINalpha"
                     /note="synonyms: PRO1489, MGC22256"
                     /db_xref="LocusID:55450"
     CDS             145. .306
                     /gene="CaMKIINalpha"
                     /codon_start=1
                     /product="CaMKIINalpha protein"
                     /protein_id="AAH20630.1"
                     /db_xref="GI:18089056"
                     /db_xref="LocusID:55450"
                     /translation="MKRESTLKFITSLLPTMKSTTCIWYQAGRQMRREEEEEEEKALG
                     SSAKIKIKK"
ORIGIN

  Query Match             31.0%;  Score 987.8;  DB 9;  Length 1377;
  Best Local Similarity   98.2%;  Pred. No. 1.3e-128;
  Matches 1024;  Conservative    0;  Mismatches    7;  Indels   12;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 67

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 127

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 187

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 247

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 307

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 367

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 427

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 487

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 547

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 607

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        608 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATAATCTGTTTTATTTTGGTCA 667

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 727

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 787

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 847

Qy       2998 ACACACACACACACACACACAC----------GAAAAACAAAGAAAAAAATGCTTGA 3046
             |||||||||||||||||||||||          ||||  ||  ||||||||||||||
Db        848 ACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGA 907

Qy       3047 GCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATAT 3106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 GCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATAT 967

Qy       3107 TATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT- 3165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        968 TATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATA 1027

Qy       3166 TCTTGAAATGCCAAAAGGCTTTT 3188
             ||||||||||||||||||| |||
Db       1028 TCTTGAAATGCCAAAAGGCATTT 1050


RESULT 9
AR454597
LOCUS       AR454597               1876 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 70 from patent US 6682888.
ACCESSION   AR454597
VERSION     AR454597.1  GI:42688066
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1876)
  AUTHORS   Loring,J.F., Tingley,D.W. and Edwards,C.M.
  TITLE     Genes expressed in alzheimer's disease
  JOURNAL   Patent: US 6682888-A 70 27-JAN-2004;
FEATURES             Location/Qualifiers
     source          1. .1876
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
  Query Match           30.2%;  Score 963.8;  DB 6;  Length 1876;
  Best Local Similarity  96.5%;  Pred. No. 2.9e-125;
  Matches 1004;  Conservative    0;  Mismatches   27;  Indels    9;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 579

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 639

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 699

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 759

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAAT-AAAAAAATAA 2456
             |||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        760 AGGAGGGGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAAATAA 819

Qy       2457 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 2516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 879

Qy       2517 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 2576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 939

Qy       2577 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 2636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 999

Qy       2637 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 2696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 1059

Qy       2697 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTAT 2756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTAT 1119

Qy       2757 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 2816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 1179

Qy       2817 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 2876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 1239

Qy       2877 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 1299

Qy       2937 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 2996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 1359

Qy       2997 -------CACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCT 3049
                      ||||||||||||||||||||||||||||
Db       1360 NGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGAAAAACAAAGAAAAAAATGCTTGAGCT 1419

Qy       3050 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 1479

Qy       3110 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCT 3168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1480 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCT 1539

Qy       3169 TGAAATGCCAAAAGGCTTTT 3188
                ||||||||||||||||| |||
Db       1540 TGAAATGCCAAAAGGCATTT 1559
```

```
RESULT 10
AL807249
LOCUS       AL807249              248479 bp    DNA     linear   ROD 24-OCT-2002
DEFINITION  Mouse DNA sequence from clone RP23-25C1 on chromosome 4, complete
            sequence.
ACCESSION   AL807249
VERSION     AL807249.8  GI:24395096
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 248479)
  AUTHORS   Blakey,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 25, 2002 this sequence version replaced gi:24366558.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-25C1 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .248479
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="4"
                     /clone="RP23-25C1"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match           25.2%;  Score 805.2;  DB 10;  Length 248479;
  Best Local Similarity  62.6%;  Pred. No. 4.8e-103;
  Matches 2076;  Conservative    0;  Mismatches  893;  Indels  345;  Gaps   39;

Qy          2 CCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGC-TCCCAGTCTGCCTCCC 60
              ||  ||||||   |    |||   |  | | ||||||||| || || |||||  |    |
Db     239212 CCCTCCCTCCTTTTTGCTCCCCTCCCCCTCGGCCGCTCGCTGCTTCACAGTCCCCGAGTC 239271

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
                |||  |||||||||||||  ||| |||||||  |||||||  ||   ||||||||||||
Db     239272 GCCGCCGCTGCCGCCGCCCGCATCCGCGGCAGGACTGGCCCGCGCGCTCCCCGGGGTGCG 239331
```

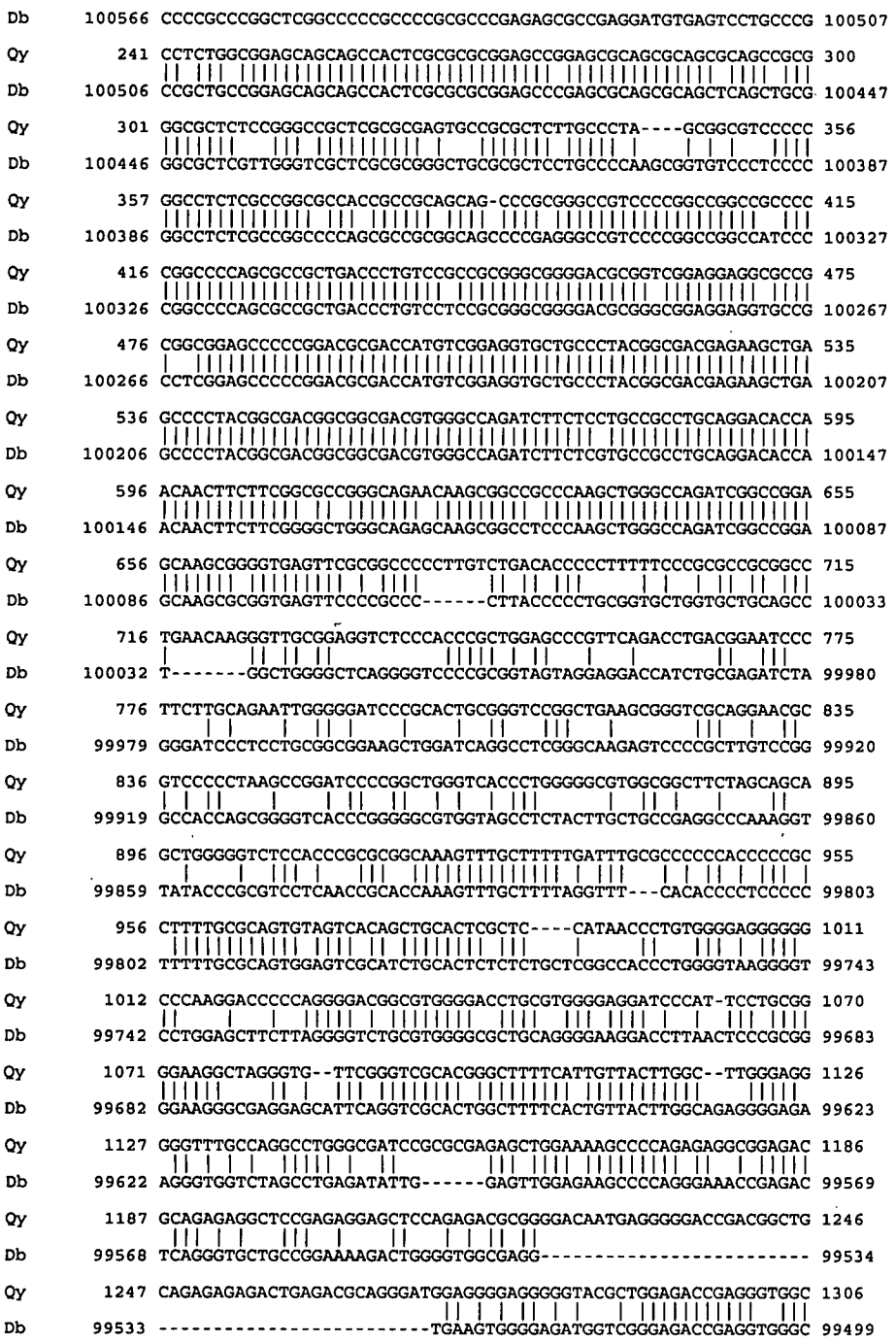

Qy 121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
Db 239332 CCCACCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACACCCTCCTCCCCGCGCGTCGCGG 239391
Qy 181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
Db 239392 CCCCGCCCGGCTCGGCCCCCGCCCCGCGCCCGAGAGCGCCGAGGATGTGAGTCCTGCCCG 239451
Qy 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Db 239452 CCGCTGCCGGAGCAGCAGCCACTCGCGCGCGGAGCCCGAGCGCAGCGCAGCTCAGCTGCG 239511
Qy 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTA----GCGGCGTCCCCC 356
Db 239512 GGCGCTCGTTGGGTCGCTCGCGCGGGCTGCGCGCTCCTGCCCCAAGCGGTGTCCCTCCCC 239571
Qy 357 GGCCTCTCGCCGGCGCCACCGCCGCAGCAG-CCCGCGGGCCGTCCCCGGCCGGCCGCCCC 415
Db 239572 GGCCTCTCGCCGGCCCCAGCGCCGCGGCAGCCCCGAGGGCCGTCCCCGGCCGGCCATCCC 239631
Qy 416 CGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCG 475
Db 239632 CGGCCCCAGCGCCGCTGACCCTGTCCTCCGCGGGCGGGGACGCGGGCGGAGGAGGTGCCG 239691
Qy 476 CGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGA 535
Db 239692 CCTCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGA 239751
Qy 536 GCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCA 595
Db 239752 GCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCGTGCCGCCTGCAGGACACCA 239811
Qy 596 ACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGA 655
Db 239812 ACAACTTCTTCGGGGCTGGGCAGAGCAAGCGGCCTCCCAAGCTGGGCCAGATCGGCCGGA 239871
Qy 656 GCAAGCGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCC 715
Db 239872 GCAAGCGCGGTGAGTTCCCCGCCC------CTTACCCCCTGCGGTGCTGGTGCTGCAGCC 239925
Qy 716 TGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCC 775
Db 239926 T-------GGCTGGGGCTCAGGGGTCCCCGCGGTAGTAGGAGGACCATCTGCGAGATCTA 239978
Qy 776 TTCTTGCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGC 835
Db 239979 GGGATCCCTCCTGCGGCGGAAGCTGGATCAGGCCTCGGGCAAGAGTCCCCGCTTGTCCGG 240038
Qy 836 GTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCA 895
Db 240039 GCCACCAGCGGGGTCACCCGGGGGCGTGGTAGCCTCTACTTGCTGCCGAGGCCCAAAGGT 240098
Qy 896 GCTGGGGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGC 955
Db 240099 TATACCCGCGTCCTCAACCGCACCAAAGTTTGCTTTTAGGTTT---CACACCCCTCCCCC 240155
Qy 956 CTTTTGCGCAGTGTAGTCACAGCTGCACTCGCTC----CATAACCCTGTGGGGAGGGGGG 1011
Db 240156 TTTTTGCGCAGTGGAGTCGCATCTGCACTCTCTCTGCTCGGCCACCCTGGGGTAAGGGGT 240215
Qy 1012 CCCAAGGACCCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATC-CCATTCCTGCGG 1070
Db 240216 CCTGGAGCTTCTTAGGGGTCTGCGTGGGGCGCTGCAGGGGAAGGACCTTAACTCCCGCGG 240275
Qy 1071 GGAAGGCTAGGGTG--TTCGGGTCGCACGGGCTTTTCATTGTTACTTGGC--TTGGGAGG 1126
Db 240276 GGAAGGGCGAGGAGCATTCAGGTCGCACTGGCTTTTCACTGTTACTTGGCAGAGGGGAGA 240335
Qy 1127 GGGTTTGCCAGGCCTGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGAC 1186
Db 240336 AGGGTGGTCTAGCCTGAGATAT------TGGAGTTGGAGAAGCCCCAGGGAAACCGAGAC 240389
Qy 1187 GCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTG 1246
Db 240390 TCAGGGTGCTGCCGGAAAAGACTGGGGTGGCGAGG---------- 240424
Qy 1247 CAGAGAGAGACTGAGACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGC 1306
Db 240425 ----------TGAAGTGGGGAGATGGTCGGGAGACCGAGGTGGGC 240459
Qy 1307 AGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACG 1366
Db 240460 TGAGAGCTA----------AAGCGAGGGAGCTGTAAGGGAGCCTGGGACT 240499
Qy 1367 CGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACA-----CCTTA 1421
Db 240500 CAAGTCTAGGGAGGACGCTGAAG-TGCTGGAACAATGAGGGTGACAGACAAGGAGACCTA 240558
Qy 1422 GACAGGGAGAGACAGAGACCTCGATCGGCTGCCGG--CCGTCGCGCCGAGGGACGATGGA 1479
Db 240559 CGGGTAGAGACCCAGAGTTCTGGACAGGTGATCTGATCATCCAGTACAGGGACCGGTGGA 240618
Qy 1480 GGGACTGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTG 1539
Db 240619 GGGACTGAGAGAGGCCAGTGGAGGCAGTGAGGCAGGCGGAAGAGGCAACTGAAGGTAAAG 240678
Qy 1540 TCCCTGGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGC 1599
Db 240679 CCCCAGGC-GGCAGCGAAGGGAGGTTCTGCCCTTTGCGGTGGCACAGAGAGCGGAGGGGC 240737
Qy 1600 CGAGTCAGTCGGCCAGCGGCTGGGAGAGGGCGCGCAGGAGGGGG-CGCCCGCCCAGGCCA 1658
Db 240738 GGAGGCTACAGACCGGGGAACGGGGAAGGGAGAACGGGAGGGGGACACCCGTTCAGGCCA 240797
Qy 1659 GGCCCTAAC-CCCCACCCGCTGCGCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTT 1717
Db 240798 GGCCTAAACTTTGCAGCCGCTGATCGGCGTGGGAACCTACCCCAGCTTTCCTTCCT-GCT 240856
Qy 1718 CCGCTCATCTCCGCACCTAGCCTTGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTC 1777
Db 240857 CCGGGCATCTTTGAACCCGGCCTGGGCTACCG----------CCACCAGTGATC 240900
Qy 1778 CCGAGACCCAAAGACTTTGGCTCCCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAAT 1837
Db 240901 TGAAGACCTCAGTTTCCCTTTCTCTCCTTAAGAATCAGCCTAGGAGACTTCTTTGTAAAG 240960
Qy 1838 TAGTGCGCCTGGTGCGGGGAGGACGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGG 1897
Db 240961 GGTTGCATCGGGTTGCAGTCG----------AG 240983
Qy 1898 CTGGGTTGGCGCCCTCTG--TTTCCTTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGG 1955
Db 240984 CTGGGATGGCTCCCTCTGGCTTGCTCCTTCTTAGAATCCATCTAGGATGAGGGCT----- 241038
Qy 1956 ATAGAGGGTGGTGGGTGGTTCAGAGGAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGG 2015
Db 241039 ----AGGCAGTGGGGCAGTTCAGAACCAGGCAGGGAAGGGAGCCTAG---GGCAATGGGG 241091
Qy 2016 ATGGAGCTGTTTCACCGCGCAGTGAGCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCT-- 2073
Db 241092 CTGTTCACCCTGCTTGGAACAGCCAGCCTTCCTTCCTCGCCCTCCCCTGCACCTCCATAA 241151
Qy 2074 ----------CCCACTCTGGGCATAACGGGAAATGTCAGAGACCTCTGGCTAGGCCCCAG 2123
Db 241152 TGGGGATAGACAAAGTATCTGCGACCCCTGGCTAGCCCAAGCTCTCTCACATTTCTCTTC 241211
Qy 2124 CGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGA 2183
Db 241212 CCCTCCCCCCCCCTTTTTTTTTTTTTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGA 241271
Qy 2184 TGACGTGCTGAAAAATATGACCGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACA 2243
Db 241272 TGACGTGCTGAAAACCATGACCGACAAGGCACCCCCTGGTGTCTAACTGCCCCAAAGACA 241331
Qy 2244 ATGAGTTAAGGGAGAGAATAAGAA-------CGGCGGTAACAGTTATTGGCAAAAAGCAT 2296

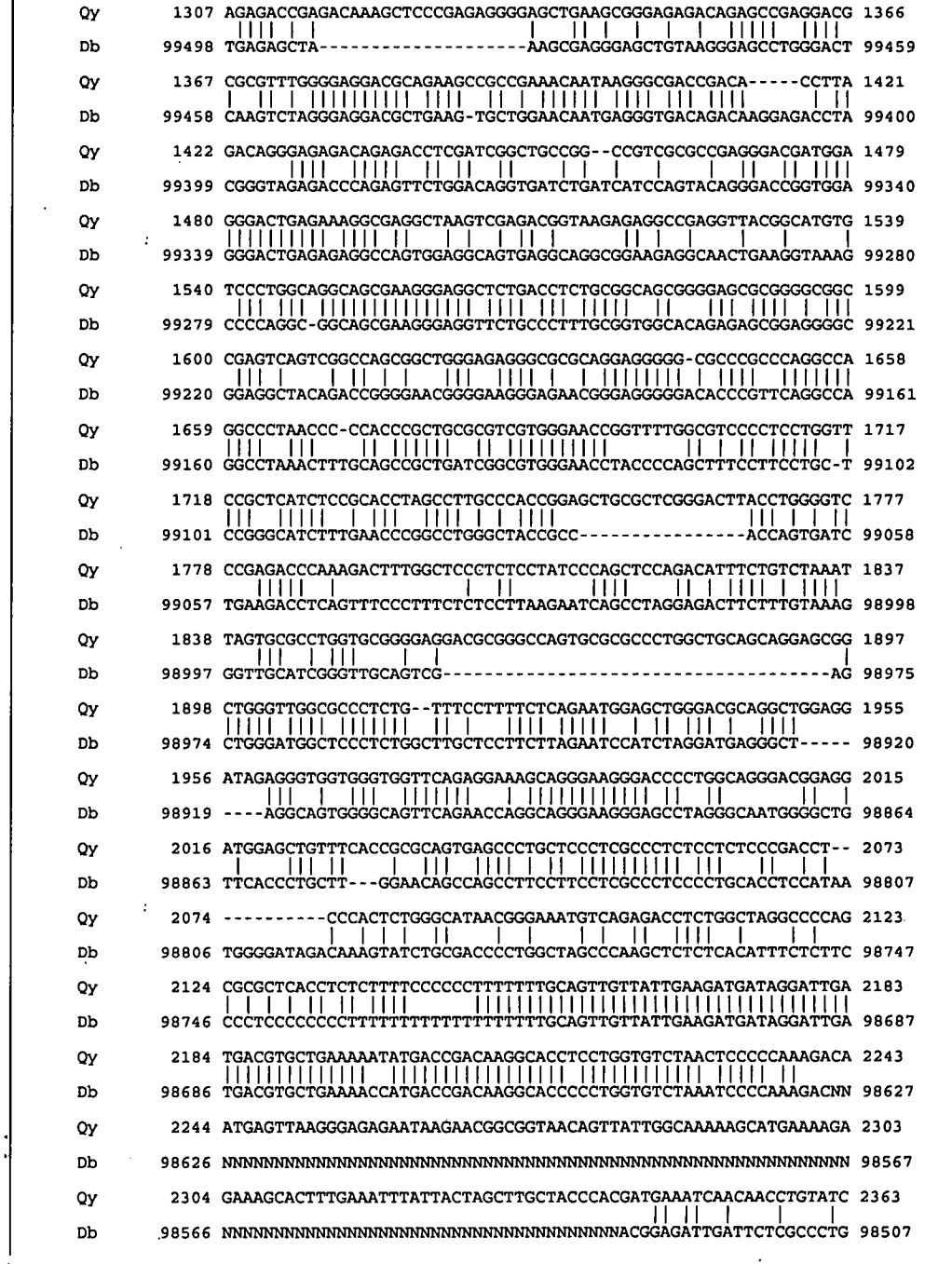


RESULT 11
AC107845/c

LOCUS AC107845 170878 bp DNA linear HTG 13-MAR-2004
DEFINITION Mus musculus clone RP23-308C7, WORKING DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION AC107845
VERSION AC107845.3 GI:45430114
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 170878)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-308C7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170878)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 170878)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 13, 2004 this sequence version replaced gi:20148025.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```
             ---------- Project Information
                 Center project name: L20609
                 Center clone name: 308_C_7
             ---------- Summary Statistics
                 Sequencing vector: Plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.960731
                 Consensus quality: 160600 bases at least Q40
                 Consensus quality: 162908 bases at least Q30
                 Consensus quality: 163778 bases at least Q20
                 Insert size: 172000; agarose-fp
                 Insert size: 164695; sum-of-contigs
                 Quality coverage: 7.7 in Q20 bases; agarose-fp
                 Quality coverage: 8.0 in Q20 bases; sum-of-contigs
             ----------.
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 24 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *        1      642: contig of 642 bp in length
             *      643      742: gap of 100 bp
             *      743    12238: contig of 11496 bp in length
             *    12239    12338: gap of 100 bp
             *    12339    26669: contig of 14331 bp in length
             *    26670    26769: gap of 100 bp
             *    26770    43832: contig of 17063 bp in length
             *    43833    43932: gap of 100 bp
             *    43933    48142: contig of 4210 bp in length
             *    48143    48242: gap of 100 bp
             *    48243    75380: contig of 27138 bp in length
             *    75381    75480: gap of 100 bp
             *    75481    76339: contig of 859 bp in length
             *    76340    76439: gap of 100 bp
             *    76440    77363: contig of 924 bp in length
             *    77364    77463: gap of 100 bp
             *    77464    80676: contig of 3213 bp in length
             *    80677    80776: gap of 100 bp
             *    80777    93584: contig of 12808 bp in length
             *    93585    93684: gap of 100 bp
             *    93685    98528: contig of 4844 bp in length
             *    98529    98628: gap of 100 bp
             *    98629   114722: contig of 16094 bp in length
             *   114723   114822: gap of 100 bp
             *   114823   127312: contig of 12490 bp in length
             *   127313   127412: gap of 100 bp
             *   127413   141549: contig of 14137 bp in length
             *   141550   141649: gap of 100 bp
             *   141650   142476: contig of 827 bp in length
             *   142477   142576: gap of 100 bp
             *   142577   144654: contig of 2078 bp in length
             *   144655   144754: gap of 100 bp
             *   144755   154484: contig of 9730 bp in length
             *   154485   154584: gap of 100 bp
             *   154585   156750: contig of 2166 bp in length
             *   156751   156850: gap of 100 bp
             *   156851   158875: contig of 2025 bp in length
             *   158876   158975: gap of 100 bp
             *   158976   160836: contig of 1861 bp in length
             *   160837   160936: gap of 100 bp
             *   160937   164666: contig of 3730 bp in length
             *   164667   164766: gap of 100 bp
             *   164767   166817: contig of 2051 bp in length
             *   166818   166917: gap of 100 bp
             *   166918   168913: contig of 1996 bp in length
             *   168914   169013: gap of 100 bp
             *   169014   170878: contig of 1865 bp in length.
FEATURES             Location/Qualifiers
     source          1. .170878
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-308C7"
                     /clone_lib="RPCI-23 Female Mouse BAC"
          misc_feature    1. .642
                     /note="assembly_fragment"
          misc_feature    743. .12238
                     /note="assembly_fragment"
          misc_feature    12339. .26669
                     /note="assembly_fragment"
          misc_feature    26770. .43832
                     /note="assembly_fragment"
          misc_feature    43933. .48142
                     /note="assembly_fragment"
          misc_feature    48243. .75380
                     /note="assembly_fragment"
          misc_feature    75481. .76339
                     /note="assembly_fragment"
          misc_feature    76440. .77363
                     /note="assembly_fragment"
          misc_feature    77464. .80676
                     /note="assembly_fragment"
          misc_feature    80777. .93584
                     /note="assembly_fragment"
          misc_feature    93685. .98528
                     /note="assembly_fragment"
          misc_feature    98629. .114722
                     /note="assembly_fragment"
          misc_feature    114823. .127312
                     /note="assembly_fragment"
          misc_feature    127413. .141549
                     /note="assembly_fragment"
          misc_feature    141650. .142476
                     /note="assembly_fragment"
          misc_feature    142577. .144654
                     /note="assembly_fragment"
          misc_feature    144755. .154484
                     /note="assembly_fragment"
          misc_feature    154585. .156750
                     /note="assembly_fragment"
          misc_feature    156851. .158875
                     /note="assembly_fragment"
          misc_feature    158976. .160836
                     /note="assembly_fragment"
          misc_feature    160937. .164666
                     /note="assembly_fragment"
          misc_feature    164767. .166817
                     /note="assembly_fragment"
          misc_feature    166918. .168913
                     /note="assembly_fragment"
          misc_feature    169014. .170878
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"

  Query Match             23.0%;  Score 734.4;  DB 2;  Length 170878;
  Best Local Similarity   59.5%;  Pred. No. 3.9e-93;
  Matches 1923;  Conservative    0;  Mismatches 1021;  Indels  286;  Gaps   35;

Qy          2 CCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGC-TCCCAGTCTGCCTCCC 60
              ||  ||||||   |    |||  |  | | ||||||||| || || |||||  |   |
Db     100746 CCCTCCCTCCTTTTTGCTCCCCTCCCCCTCGGCCGCTCGCTGCTTCACAGTCCCCGAGTC 100687

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
                |||  |||||||||||||  ||| |||||||  |||||||  ||   ||||||||||
Db     100686 GCCGCCGCTGCCGCCGCCCGCATCCGCGGCAGGACTGGCCCGCGCGCTCCCCGGGGTGCG 100627

Qy        121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
              ||| ||||||||||||||||||||||||||||||||  |   |  || | ||  ||||||
Db     100626 CCCACCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACACCCTCCTCCCCGCGCGTCGCGG 100567

Qy        181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
              |||||||||||  |||||||||||| | |||| || |||||||||||||||||||||| ||
```

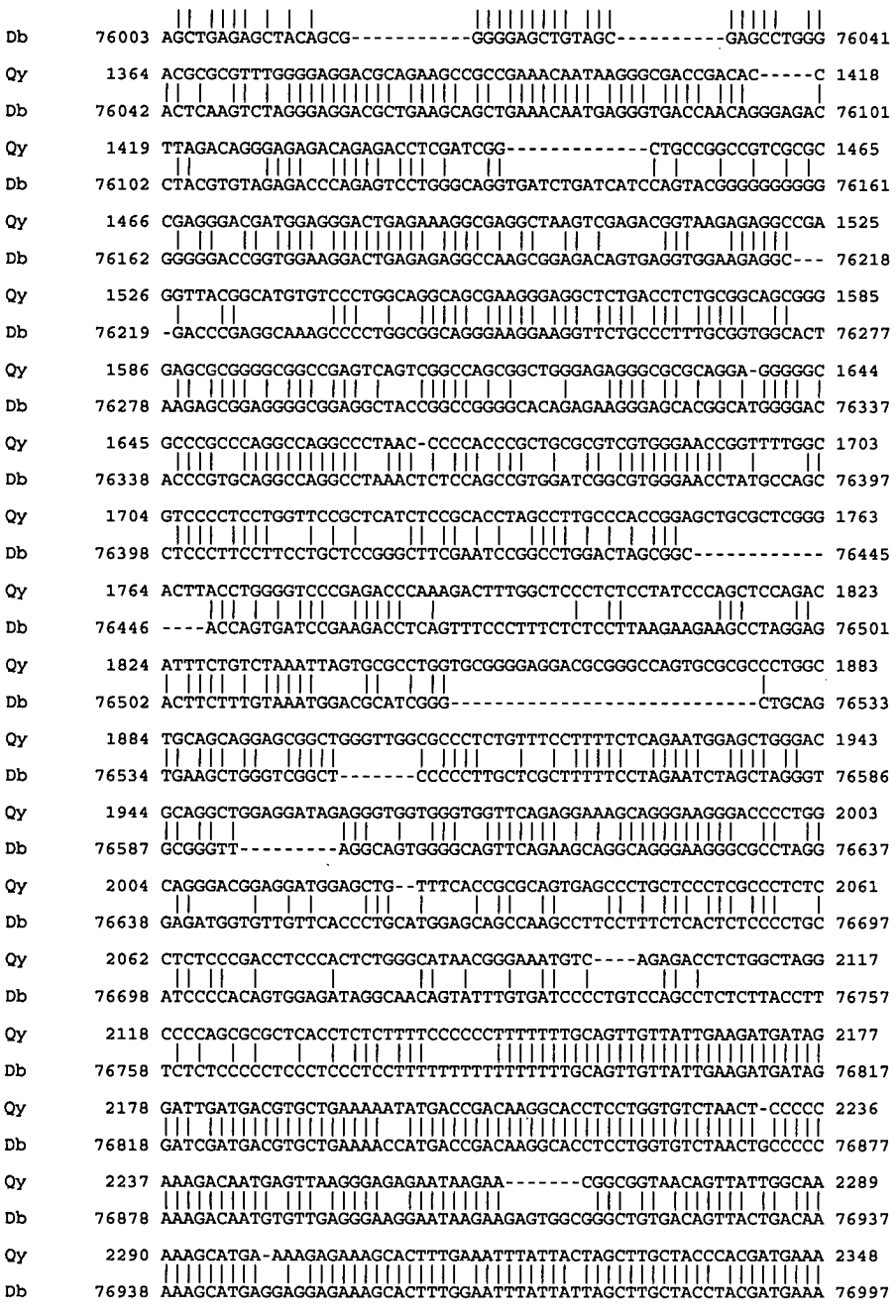


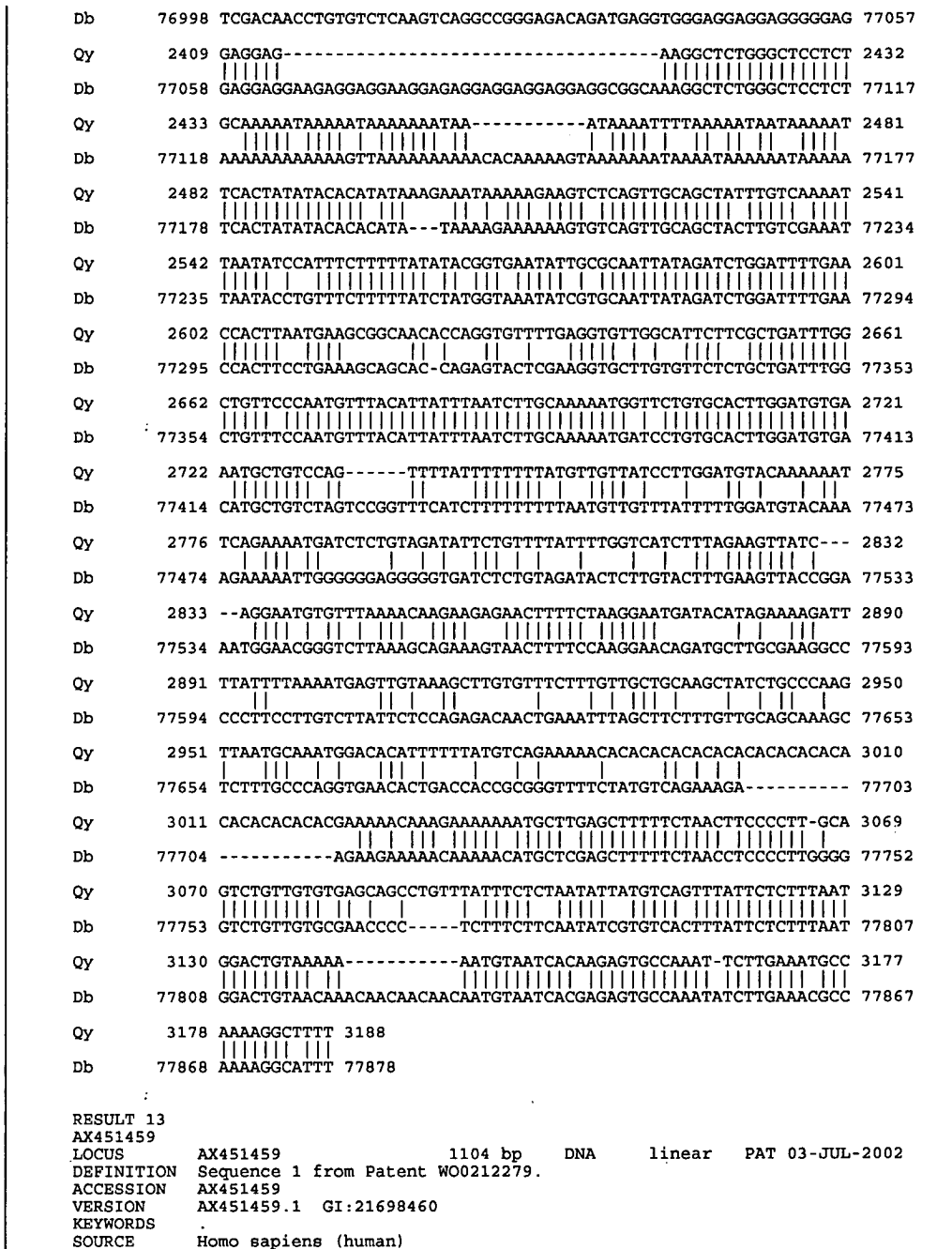


```
Qy       2364 TGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGGAGGAGGAGGAGGAAGGCTCTG 2423
             ||||   |  ||| |   |                                  | |
Db      98506 TGGTGGAATTCCGTGGCTC-------------------------TGGGCTCCTCTACAAA 98472

Qy       2424 GGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTC 2483
               ||    | ||||| || |   ||||| |||| |||      |||| ||  |||| ||
Db      98471 AACTAACGTTAAAAAAAAACACACAAAAAGTAAAAAAATAAAATAAAAAAAATAAAAATC 98412

Qy       2484 ACTATATACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTA 2543
             |||||||||||| ||| | |   | ||| |||||||||||||||||||||||||| ||||
Db      98411 ACTATATACACACATATAAA---AGAAAAAAGTCTCAGTTGCAGCTATTTGTCAAGATTA 98355

Qy       2544 ATATCCATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACC 2603
             ||| |  |||||||||||||| ||| |||||| ||||||||| |||||| ||||||||
Db      98354 ATACCTGTTTCTTTTTATATATGGTAAATATTGTGCAATTATAGCTCTGGAGTTTGAACC 98295

Qy       2604 ACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCT 2663
             ||||   |||   ||| ||||       |  ||||| | || ||||  ||||||||||||
Db      98294 ACTTCCCGAA--AGCAGCACCGACCACGTGAAGGTGCTTGCCTTCTCTGCTGATTTGGCT 98237

Qy       2664 GTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAA 2723
             |||||||||||||||| |||| ||||          || |  |    |  ||  ||
Db      98236 GTTCCCAATGTTTACATAATTTCATCT----------GTGCACTTGGATGTGACATGCGG 98187

Qy       2724 TGCTGTCCAGTTTTATTTTTTTTATGTTG-TTATCCTTGGATGTACAAAAAATTCAGAAA 2782
             |   |||||||||| |||||||| |||||| |||| |||||||||||||||||| |
Db      98186 TCTAGTCCAGTTTTTTTTTTTTTAATGTTGTTTATTCTTGGATGTACAAAAGAAAATTGGG 98127

Qy       2783 ATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGT 2842
               |      |      | |   |  | |||  | |   |     ||    | |     |
Db      98126 GGCAGGGGAATGACCTCTGTAGATACTGTTGTACTTGGGTCATCTTTGAATGAACCCCAT 98067

Qy       2843 TTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAG------------AT 2889
             || |||  ||||  |||||||||| ||| |||| | | |  |||             |
Db      98066 TTTAAAGCAGAAAGTAACTTTTCTACGGACTGATGCTTTGGGAAGGCCCCTTCCTGGTCT 98007

Qy       2890 TTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAA 2949
             | | ||  | |   |  || |||||||| | |||||||||| | |       |  || |
Db      98006 TGTTTTCGAGAGACAACTGAAAAGCTTGCGCTTCTTTGTTGCAGTACAGCTCTTTGCCCA 97947

Qy       2950 GTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACACACACACAC 3009
             | | |      |         |||| ||||||||||
Db      97946 GGTTAACACCAACAGTGGGTTTTATATGTCAGAA------------------------- 97913

Qy       3010 ACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCA 3069
               | |   | | |||||||||||| |||  |||||||||||||||| | ||||   |
Db      97912 --AGAAGAAGAAGAAAAACAAAGACAAACTTGCTTGAGCTTTTTCTAGCCTCCCTGGGGG 97855

Qy       3070 GTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTT----ATTCTCTT 3125
              |||||| |||  | ||   ||  ||||  |||||| ||||| |||    ||||||||
Db      97854 CTCTGTTGAGTGA--ACCCCTCTTTCTTCTTCAATATTGTGTCACTTTATCCATTCTCTT 97797

Qy       3126 TAATGGACTGTAAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATG 3175
              |||||||||||||||||    |  |  | |   | ||  |  | | | |
Db      97796 TAATGGACTGTAAAAAAAAAACAAAAACAAACAACAAATGTAATCACAAG 97747
```

```
RESULT 12
AC109006
LOCUS       AC109006               259224 bp    DNA     linear   HTG 26-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-44I19, *** SEQUENCING IN PROGRESS
            ***, 13 unordered pieces.
ACCESSION   AC109006
VERSION     AC109006.5  GI:23321630
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 259224)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 259224)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 259224)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 26, 2002 this sequence version replaced gi:22953899.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
```

```
            Contact: hgsc-help@bcm.tmc.edu
          ---------- Project Information
            Center project name: GPPS
            Center clone name: CH230-44I19
          ---------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 162075 bases at least Q40
            Consensus quality: 169729 bases at least Q30
            Consensus quality: 175172 bases at least Q20
            Estimated insert size: 196813; sum-of-contigs estimation
            Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
          ----------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 13 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1    19653: contig of 19653 bp in length
          *    19654    19753: gap of unknown length
          *    19754   103202: contig of 83449 bp in length
          *   103203   103302: gap of unknown length
          *   103303   151034: contig of 47732 bp in length
          *   151035   151134: gap of unknown length
          *   151135   173513: contig of 22379 bp in length
          *   173514   173613: gap of unknown length
          *   173614   228828: contig of 55215 bp in length
          *   228829   228928: gap of unknown length
          *   228929   245283: contig of 16355 bp in length
          *   245284   245383: gap of unknown length
          *   245384   246664: contig of 1281 bp in length
          *   246665   246764: gap of unknown length
          *   246765   247782: contig of 1018 bp in length
          *   247783   247882: gap of unknown length
          *   247883   249267: contig of 1385 bp in length
          *   249268   249367: gap of unknown length
          *   249368   251556: contig of 2189 bp in length
          *   251557   251656: gap of unknown length
          *   251657   253199: contig of 1543 bp in length
          *   253200   253299: gap of unknown length
          *   253300   254502: contig of 1203 bp in length
          *   254503   254602: gap of unknown length
          *   254603   259224: contig of 4622 bp in length.
FEATURES             Location/Qualifiers
     source          1. .259224
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-44I19"
     misc_feature    1. .1160
                     /note="wgs_contig"
     misc_feature    18526. .19653
                     /note="wgs_contig"
     misc_feature    33930. .35449
                     /note="wgs_contig"
     misc_feature    119013. .122247
                     /note="wgs_contig"
     misc_feature    212187. .213772
                     /note="wgs_contig"
ORIGIN

  Query Match             22.4%;  Score 713.4;  DB 2;  Length 259224;
  Best Local Similarity   63.1%;  Pred. No. 3.4e-90;
  Matches 1975;  Conservative    0;  Mismatches  811;  Indels  345;  Gaps   43;

Qy        196 CCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCGCCTCTGGCGGAGCAG 255
              ||||||||| | |||| || |||||||||||||||||| ||| |||  ||| ||||||||
Db      74955 CCCCCGCCCCGCGCCCGAGAGCGCCGAGGATGTGAGTCGTGCCCGC--CTGCCGGAGCAG 75012

Qy        256 CAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCGGGCGCTCTCCGGGCC 315
              ||||||||||||||||||| |||||||||||| |||| |||||||||    ||| |
Db      75013 CAGCCACTCGCGCGCGGAGCCCGAGCGCAGCGCAGCTCAGCTGCGGGCGCTCGTTGGGTC 75072

Qy        316 GCTCGCGCGAGTGCCGCGCTCTTGCCCTA----GCGGCGTCCCCCGGCCTCTCGCCGGCG 371
              ||||||||| |   ||||||| ||||| |    | | |   |||||||| |||||||||
Db      75073 GCTCGCGCGGGCTGCGCGCTCCTGCCCCAAGCGGTGTCTCTCCCCGGCCACTCGCCGGCC 75132

Qy        372 CCACCGCCGCAGCAG-CCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCCCCAGCGCCGC 430
              ||| | |||| |||| |||| ||||||||||||||||||||   ||||||||||||||||
Db      75133 CCAGCACCGCGGCAGCCCCGAGGGCCGTCCCCGGCCGGCCATCCCCGGCCCCAGCGCCGC 75192

Qy        431 TGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCGGAGCCCCCGG 490
              |||||||||| |||||||||||||||||| |||||||| |||||  ||||||||||||
Db      75193 TGACCCTGTCCTCCGCGGGCGGGGACGCGGGCGGAGGAGGTGCCGCCTCGGAGCCCCCGG 75252

Qy        491 ACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCCTACGGCGACG 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75253 ACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCCTACGGCGACG 75312

Qy        551 GCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAACTTCTTCGGCG 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75313 GCGGCGACGTGGGCCAGATCTTCTCGTGCCGCCTGCAGGACACCAACAACTTCTTCGGCG 75372

Qy        611 CCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAGCGGGGTGAGT 670
              |||||||| |||||||||||||||||||||||||||||||||||||||||| |||||||
Db      75373 CCGGGCAGAGCAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAGCGCGGTGAGT 75432

Qy        671 TCGCGGCCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAACAAGGGTTGCG 730
              || | ||||      || || ||||    | |  | || || |||||      ||  |||
Db      75433 TCCCCGCCC------CTTACCCCCCGCGGTGCTGGTGCTGCAGCCTG----GCGGGGGCG 75482

Qy        731 GAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTTGCAGAATTGG 790
              ||| |||| | |  || |   |                          ||| | ||
Db      75483 CAGGGGTCCCCGCGGTGGGGGGTGG----------------------CATCTGCGGGATCTA 75522

Qy        791 GGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCCCCTAAGCCGG 850
              |||||||| | ||||||     |||| | | || | | || | | |||  | |  | | |
Db      75523 GGGATCCCTC-CTGCGGCGGAAGCTGGATCAGGCCTCGGGCAAGAGTCTGCGTTTGTCTG 75581

Qy        851 ATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGGGGGTC----- 905
              || |   ||||    || ||||||||||  |||||||   |||||||  ||  |
Db      75582 GGCCACCAGTGGGGTCACCCGGGGGCGTGGTAGCTTCTACTTGCAGCTGAGGCCCAAAGG 75641

Qy        906 ----------TCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCG 954
                       ||| |   |||  |||||||||||||| |||||    | | || | |||
Db      75642 TTATACCCACGTCCTCAACCGCACCAAAGTTTGCTTTTAGATTT---CACACCCCTCCCC 75698

Qy        955 CCTTTTGCGCAGTGTAGTCACAGCTGCACTC-----GCTCCATAACCCTGTGGGGAGGGG 1009
              | |||||||||||| |||| || ||| ||||     ||||   ||||||  |  |||||
Db      75699 CTTTTTGCGCAGTGGAGTCGCATCTGTACTCTCTCTGCTCTGCCACCCTGGAGTAAGGGG 75758

Qy       1010 GGCCCAAGGACCCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCAT-TCCTGC 1068
              ||  ||   |  ||| | | | |  |||| ||| |||| | |  ||| |
Db      75759 TCCCGGAGCTTCTTAGGGGTCTGTGGGAGTGGGGTGCGGGGGAAGGACCTTAACTCCCGT 75818

Qy       1069 GGGGAAGGCTAGG-GTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTT---GGGA 1124
              ||||||||| ||| | |||  |||||||| || |||| | |||||  | ||    ||||
Db      75819 GGGGAAGGCGAGGAGCGTTGAGGTCGCACAGGTTTTTGACTGTTAACTTGCCAGAGGGGA 75878

Qy       1125 GGGGGTTTGCCAGGCCTGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAG 1184
              || | |    |||||||  ||     |     |||| ||||||||||||  ||||
Db      75879 AAAGGGTGGTGTAGCCTGGGATATTGATGT------TGGAGAAGCCCCAGAGAAACGGAC 75932

Qy       1185 ACGCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGC 1244
              |||||||                                                  ||
Db      75933 ACGCAGAT----------------------------------------------------GC 75942

Qy       1245 TGCAGAGAGAGACTGAGACGCAGGGATGGAGGGGAG-GGGGTACGCTGGAGACCGAGGGT 1303
              |||    | |||||| |   | | | || ||||  | | | ||| | || ||||||||
Db      75943 TGCCCTAAAAGACTGGGGTACCGAGGTGAAGGGTGGAGAGATACTCAGGGGACCGAGGTG 76002

Qy       1304 GGCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGG 1363
```





```
RESULT 13
AX451459
LOCUS       AX451459               1104 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0212279.
ACCESSION   AX451459
VERSION     AX451459.1  GI:21698460
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bihain,B., Bour,B. and Bougueleret,L.
  TITLE     Schizophrenia related gene and protein
  JOURNAL   Patent: WO 0212279-A 1 14-FEB-2002;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .1104
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             87. .344
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD37813.1"
                     /db_xref="GI:21698461"
                     /translation="MWEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRP
                     PKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPLVSNSPKTMS"
ORIGIN

  Query Match           19.0%;  Score 607.2;  DB 6;  Length 1104;
  Best Local Similarity  89.2%;  Pred. No. 2.5e-75;
  Matches  787;  Conservative    0;  Mismatches   63;  Indels   32;  Gaps   11;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 311

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              |||||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db        312 -CTGGTGTCTAACT-CCCCAAAGACAATGAGTTAAGGGAGAGAATAGGAACGGCGGTAAC 369

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        370 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TAC 428

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              |||||||||||||||||||||||||||||| | ||| |||||||||| |||||| |||
Db        429 CCACGATGAAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGGCGA-AGG 486

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              |||| |||   ||| |||| |   ||||| ||||| |||| | || || |||||||| |
Db        487 AGGAAGAGAGAGAGAAGAAAGGCTTGGGC-CCTCTACAAATAAAATAAAAAAAAAAATT 545

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
              ||||||  |||||     || |                | | ||  | |||  ||||
Db        546 TAAAATAATAAAATCCCTATA----------------TCCCATATAAGAATAAAAGAGTC 589

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
              |||||    ||  | |    ||| | | ||||||||||||||||    ||    ||
Db        590 TCAGT----GCAGTATTGGCAAAATTAAATCCATTTCTTTTTAATACGGG----AATATT 641

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
              |  |||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        642 GGCATTATAGATCTGGATTTTG-ACCACTTAATGAAGCGGC-ACCCCAGGTGTTTTGAGG 699

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 759

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 819

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 879

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 939

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 999

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 2997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1059

Qy       2998 ACACACACACACACACACACACACGAAAAACAAAGAAAAAAA 3039
              |||||||||||||||||||||||||||||| ||| |||||||
Db       1060 ACACACACACACACACACACACACGAAAAAAAAAAAAAAAAA 1101


RESULT 14
AY204901
LOCUS       AY204901                860 bp    mRNA    linear   PRI 20-MAY-2003
DEFINITION  Homo sapiens calcium/calmodulin-dependent protein kinase II
            inhibitor alpha (CaMKIINalpha) mRNA, complete cds.
ACCESSION   AY204901
VERSION     AY204901.1  GI:30959105
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Li,N. and Cao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2002) Institute of Immunology, Second Military
            Medical University, 800 Xiangyin Road, Shanghai, Shanghai 200433,
            China
FEATURES             Location/Qualifiers
     source          1. .860
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
     gene            1. .860
                     /gene="CaMKIINalpha"
                     /note="CaMKIINa"
     CDS             401. .637
                     /gene="CaMKIINalpha"
                     /codon_start=1
                     /product="calcium/calmodulin-dependent protein kinase II
                     inhibitor alpha"
                     /protein_id="AAO49802.1"
                     /db_xref="GI:30959106"
                     /translation="MSEVLPYGDEKLSPYGDGGDVGQIFSCRLQDTNNFFGAGQNKRP
                     PKLGQIGRSKRVVIEDDRIDDVLKNMTDKAPPGV"
ORIGIN

  Query Match           17.7%;  Score 564.2;  DB 9;  Length 860;
  Best Local Similarity  99.5%;  Pred. No. 2.6e-69;
  Matches  566;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         99 CCCGCCCGGCTCCCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACG 158
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCGCCCGGCACCCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACG 60

Qy        159 TCTCCTCCCTCCCGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCG 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTCCTCCCTCCCGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCG 120

Qy        219 CCGAGGATGTGAGTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGG 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGAGGATGTGAGTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGG 180

Qy        279 AGCGCAGCGCAGCGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTT 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCGCAGCGCAGCGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTT 240

Qy        339 GCCCTAGCGGCGTCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGT 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCCTAGCGGCGTCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGT 300
```

```
Qy        399 CCCCGGCCGGCCGCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGC 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCGGCCGGCCGCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGC 360

Qy        459 GGTCGGAGGAGGCGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTA 518
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCGGAGGAGGCGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTA 420

Qy        519 CGGCGACGAGAAGCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTG 578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGCGACGAGAAGCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTG 480

Qy        579 CCGCCTGCAGGACACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCT 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGCCTGCAGGACACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCT 540

Qy        639 GGGCCAGATCGGCCGGAGCAAGCGGGGTG 667
              ||||||||||||||||||||||||| ||
Db        541 GGGCCAGATCGGCCGGAGCAAGCGGGTTG 569


RESULT 15
HSA324422/c
LOCUS       HSA324422               616 bp    DNA     linear   PRI 18-JUL-2002
DEFINITION  Homo sapiens genomic sequence surrounding NotI site, clone
            NR1-AB5C.
ACCESSION   AJ324422
VERSION     AJ324422.1  GI:15868801
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
            Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
            Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I.,
            Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
  TITLE     NotI flanking sequences: a tool for gene discovery and verification
            of the human genome
  JOURNAL   Nucleic Acids Res. 30 (14), 3163-3170 (2002)
  MEDLINE   22131767
   PUBMED   12136098
REFERENCE   2  (bases 1 to 616)
  AUTHORS   Zabarovsky,E.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
            Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="NR1-AB5C"
ORIGIN

  Query Match             16.9%;  Score 538.8;  DB 9;  Length 616;
  Best Local Similarity   94.3%;  Pred. No. 9.2e-66;
  Matches  581;  Conservative    0;  Mismatches   32;  Indels    3;  Gaps    2;

Qy         17 CGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCCCGAACCGGCGCCGCCG 76
              || |||| ||| |||||||| | |||| |||||   |  |  |||  ||||| |||||||
Db        614 CGGCCCTTGCAACCCCGCCGTTGGCAGTTCCCA--TTTGCCTCCCAGACCGGGGCCGCCG 557

Qy         77 CCCGCACTCGCCGCAGGACCGGCCCGCCCGGC-TCCCGGGGTGCGCCCTCCTCGGTCCCG 135
              || |||||||| ||||||| |||| ||||||| ||| |||||||||| || |||||||||
Db        556 CCAGCACTCGCAGCAGGACTGGCCGGCCCGGCTTCCGGGGGTGCGCCTTCTTCGGTCCCG 497

Qy        136 CGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGGCCCCGCCCGGCCCGG 195
              |||||||||||||||||||||||||| |||||||||||| ||||| |||||||||| |
Db        496 CGCCCTCCGGGCTCGCAGGGACGTCTCTTCCCTCCCGGCTTGCGGCTCCGCCCGGCCGTG 437

Qy        196 CCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCGCCTCTGGCGGAGCAG 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCGCCTCTGGCGGAGCAG 377

Qy        256 CAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCGGGCGCTCTCCGGGCC 315
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        376 CAGCCACTCGCGCGCGGATCCGGAGCGCAGCGCAGCGCAGCCGCGGGCGCTCTCCGGGCC 317

Qy        316 GCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCCTCTCGCCGGCGCCAC 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCCTCTCGCCGGCGCCAC 257

Qy        376 CGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCCCCAGCGCCGCTGACC 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCCCCAGCGCCGCTGACC 197

Qy        436 CTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCGGAGCCCCCGGACGCG 495
              |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||
Db        196 CTGTCCGCCGCGGGCGGGGACGCGGGCGGAGGAGGCTCCGCGGCGGAGCCCCCGGACGCG 137

Qy        496 ACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCCTACGGCGACGGCGGC 555
              |||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db        136 ACCATGTCGGAGGTGCTGCCTTACGGCGACGAGAAGCTGAGCCCCTTCGGCGACGGCGGC 77

Qy        556 GACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAACTTCTTCGGCGCCGGG 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAACTTCTTCGGCGCCGGG 17

Qy        616 CAGAACAAGCGGCCGC 631
              |||| ||||||||||||
Db         16 CAGAGCAAGCGGCCGC 1


Search completed: March 28, 2005, 03:01:32
Job time : 13332.6 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 27, 2005, 20:59:17 ; Search time 1543.62 Seconds
(without alignments)
12229.764 Million cell updates/sec

Title: US-10-071-645-3
Perfect score: 3189
Sequence: 1 ccctcccctcccccctccgcc.....gaaatgccaaaaggctttta 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3189	100.0	3189	6	AAD31053	Aad31053 Human PAP
	2	1006.4	31.6	1850	11	ACN90200	Acn90200 Breast ca
	3	1002.4	31.4	1510	2	AAX20426	Aax20426 Human sec
	4	1002.4	31.4	1510	10	ADD90210	Add90210 Novel hum
	5	1002.4	31.4	1510	10	ADG90029	Adg90029 Human cDN
c	6	1002.4	31.4	2058	12	ADQ24238	Adq24238 Human sof
c	7	1000.8	31.4	1321	12	ADJ75198	Adj75198 Marker ge
c	8	1000.8	31.4	1321	12	ADN04817	Adn04817 Antipsori
	9	1000.8	31.4	1321	13	ADR25935	Adr25935 Breast ca
	10	998.4	31.3	1581	4	AAC87358	Aac87358 Human dev
	11	963.8	30.2	1876	12	ADI61702	Adi61702 Human cDN
	12	791.4	24.8	873	6	ABQ60975	Abq60975 PRO1489 p
	13	607.2	19.0	1104	6	AAD31052	Aad31052 Human PAP
c	14	583.8	18.3	869	6	ABQ43206	Abq43206 Oligonucl
	15	583.8	18.3	869	6	ABQ43207	Abq43207 Oligonucl
	16	519	16.3	620	11	ACN87560	Acn87560 Breast ca
	17	507	15.9	524	3	AAC06588	Aac06588 Human sec
c	18	499.6	15.7	674	6	ABQ43205	Abq43205 Oligonucl
	19	499.6	15.7	674	6	ABQ43204	Abq43204 Oligonucl
c	20	488.2	15.3	869	6	ABQ43209	Abq43209 Oligonucl
	21	488.2	15.3	869	6	ABQ43208	Abq43208 Oligonucl
c	22	414.2	13.0	549	10	ABT40422	Abt40422 Toxicity
c	23	414.2	13.0	549	12	ADP71513	Adp71513 Renal tox
c	24	413.4	13.0	417	5	AAF68540	Aaf68540 Human lun
c	25	413.4	13.0	417	6	ABK38451	Abk38451 cDNA enco
c	26	413.4	13.0	417	8	ACA10780	Aca10780 Human lun
c	27	413.4	13.0	417	8	ABX99731	Abx99731 Lung canc
c	28	413.4	13.0	417	10	ADH45994	Adh45994 Human lun
c	29	413.4	13.0	417	12	ADE71746	Ade71746 Human lun
c	30	413.4	13.0	417	13	ADJ19913	Adj19913 Human lun
	31	399.2	12.5	417	12	ADP18899	Adp18899 Human sec
	32	397.6	12.5	417	2	AAX97632	Aax97632 Extended
c	33	394.4	12.4	453	6	ABV97214	Abv97214 Human pan
c	34	391.4	12.3	451	11	ADT96448	Adt96448 Colon can
	35	359.2	11.3	411	6	ABL62340	Abl62340 Colon ade
	36	359.2	11.3	411	6	ABL65003	Abl65003 Lung canc
	37	359.2	11.3	411	6	ABL67288	Abl67288 Thyroid c
	38	359.2	11.3	411	6	ABL66839	Abl66839 Lung canc
c	39	328.2	10.3	674	6	ABQ43202	Abq43202 Oligonucl
	40	328.2	10.3	674	6	ABQ43203	Abq43203 Oligonucl
	41	319	10.0	336	4	AAD23580	Aad23580 Human lun
	42	319	10.0	336	10	ADD66868	Add66868 Human lun
	43	319	10.0	336	10	ADE88122	Ade88122 Human lun
	44	296.6	9.3	329	2	AAQ60934	Aaq60934 Human bra
	45	290.2	9.1	553	8	ABX97604	Abx97604 cDNA enco

ALIGNMENTS

RESULT 1
AAD31053
ID AAD31053 standard; DNA; 3189 BP.
XX
AC AAD31053;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human PAPAP gene.
XX
KW Human; PAPAP protein; schizophrenia candidate gene; g34872 gene;
KW schizophrenia; bipolar disorder; central nervous system disorder;
KW psychotic disorder; mood disorder; autism; mental retardation;
KW psychiatric disease; anxiety disorder; impulse-control disorder;
KW eating disorder; cognitive disorder; personality disorder; vaccine;
KW chromosome 1p35-p36; neuroleptic; antialcoholic; tranquiliser;
KW antidepressant; nootropic; antiaddictive; ds.
XX
OS Homo sapiens.
XX
PN WO200212279-A2.
XX
PD 14-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-IB001891.
XX
PR 07-AUG-2000; 2000US-0223482P.
XX
PA (GEST) GENSET.
XX
PI Bihain B, Bour B, Bougueleret L;
XX
DR WPI; 2002-241732/29.
XX
PT Novel isolated and purified or recombinant polynucleotide encoding PAPAP
PT protein, useful for diagnosing and treating schizophrenia, bipolar
PT disorder and other central nervous system disorders.
XX
PS Disclosure; Page 93-94; 96pp; English.
XX
CC The invention relates to human PAPAP polypeptides and polynucleotides.
CC The invention also concerns the interaction of PAPAP with schizophrenia
CC candidate gene g34872. PAPAP polypeptides, gene and anti-PAPAP antibodies

CC are useful for treating schizophrenia, bipolar disorder or related
CC central nervous system (CNS) disorders e.g. psychotic disorders, mood
CC disorders, autism, substance dependence and alcoholism, mental
CC retardation and other psychiatric diseases including cognitive, anxiety,
CC eating, impulse-control and personality disorders. They are also used in
CC vaccines. The present sequence is human PAPAP gene located on chromosome
CC 1p35-p36
XX
SQ Sequence 3189 BP; 691 A; 911 C; 959 G; 628 T; 0 U; 0 Other;

Query Match 100.0%; Score 3189; DB 6; Length 3189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120

Qy        121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180

Qy        181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240

Qy        241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300

Qy        301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360

Qy        361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420

Qy        421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480

Qy        481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540

Qy        541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600

Qy        601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660

Qy        661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720

Qy        721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780

Qy        781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840

Qy        841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900

Qy        901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960

Qy        961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020

Qy       1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080

Qy       1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140

Qy       1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200

Qy       1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260

Qy       1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320

Qy       1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380

Qy       1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440

Qy       1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500

Qy       1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560

Qy       1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620

Qy       1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680

Qy       1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740

Qy       1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800

Qy       1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860

Qy       1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920

Qy       1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
```

```
Qy       1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040

Qy       2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100

Qy       2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160

Qy       2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCTCCT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCTCCT 2220

Qy       2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280

Qy       2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340

Qy       2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400

Qy       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460

Qy       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520

Qy       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580

Qy       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640

Qy       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700

Qy       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760

Qy       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820

Qy       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880

Qy       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940

Qy       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000

Qy       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060

Qy       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120

Qy       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180

Qy       3181 AGGCTTTTA 3189
             |||||||||
Db       3181 AGGCTTTTA 3189
```

RESULT 2
ACN90200
ID ACN90200 standard; DNA; 1850 BP.
XX
AC ACN90200;
XX
DT 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 11350.
XX
KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS Homo sapiens.
XX
PN US2003099974-A1.
XX
PD 29-MAY-2003.
XX
PF 18-JUL-2002; 2002US-00198846.
XX
PR 18-JUL-2001; 2001US-0306220P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 11350; 36pp; English.
XX
CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 1850 BP; 595 A; 349 C; 430 G; 461 T; 0 U; 15 Other;

```
  Query Match             31.6%;  Score 1006.4;  DB 11;  Length 1850;
  Best Local Similarity   99.6%;  Pred. No. 9.6e-168;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 326

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 386
```

```
Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 446

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 506

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 566

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 626

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 686

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 746

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 806

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 866

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 926

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 986

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 1046

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA--ACACACACAC 2995
             ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db       1047 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1106

Qy       2996 ACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCT 3055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCT 1166

Qy       3056 AACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGT 3115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGT 1226

Qy       3116 TTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAAT 3174
             ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1227 TTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAAT 1286

Qy       3175 GCCAAAAGGCTTTT 3188
             |||||||||| |||
Db       1287 GCCAAAAGGCATTT 1300
```

RESULT 3
AAX20426
ID AAX20426 standard; DNA; 1510 BP.
XX
AC AAX20426;
XX
DT 04-MAY-1999 (first entry)
XX
DE Human secreted protein gene 15.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9906423-A1.
XX
PD 11-FEB-1999.
XX
PF 29-JUL-1998; 98WO-US015949.
XX
PR 30-JUL-1997; 97US-0054209P.
PR 30-JUL-1997; 97US-0054211P.
PR 30-JUL-1997; 97US-0054212P.
PR 30-JUL-1997; 97US-0054213P.
PR 30-JUL-1997; 97US-0054214P.
PR 30-JUL-1997; 97US-0054215P.
PR 30-JUL-1997; 97US-0054217P.
PR 30-JUL-1997; 97US-0054218P.
PR 30-JUL-1997; 97US-0054234P.
PR 30-JUL-1997; 97US-0054236P.
PR 18-AUG-1997; 97US-0055968P.
PR 18-AUG-1997; 97US-0055969P.
PR 18-AUG-1997; 97US-0055972P.
PR 19-AUG-1997; 97US-0056534P.
PR 19-AUG-1997; 97US-0056543P.
PR 19-AUG-1997; 97US-0056554P.
PR 19-AUG-1997; 97US-0056561P.
PR 19-AUG-1997; 97US-0056727P.
PR 19-AUG-1997; 97US-0056729P.
PR 19-AUG-1997; 97US-0056730P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
DR WPI; 1999-153691/13.
DR P-PSDB; AAY00272.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 1; Page 217; 312pp; English.
XX
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number is given in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. AAX20403) for increasing the stability of
CC the fused protein as compared to the human protein only. The invention
CC relates to 83 novel genes and their fragments (nucleic acid sequences:
CC AAX20412-X20499; amino acid sequences AAY00258-Y00377) which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Also, pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see AAX20412 for
CC described uses)
XX
SQ Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;

Query Match 31.4%; Score 1002.4; DB 2; Length 1510;
Best Local Similarity 99.2%; Pred. No. 4.6e-167;

```
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db         892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         952 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db        1132 AAATGCCAAAAGGCATTT 1149
```

```
RESULT 4
ADD90210
ID   ADD90210 standard; cDNA; 1510 BP.
XX
AC   ADD90210;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Novel human secreted protein cDNA seq id 25.
XX
KW   gene therapy; cytostatic; cancer; human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003199683-A1.
XX
PD   23-OCT-2003.
XX
PF   30-MAR-2001; 2001US-00820649.
XX
PR   30-JUL-1997;   97US-0054209P.
PR   30-JUL-1997;   97US-0054211P.
PR   30-JUL-1997;   97US-0054212P.
PR   30-JUL-1997;   97US-0054213P.
PR   30-JUL-1997;   97US-0054214P.
PR   30-JUL-1997;   97US-0054215P.
PR   30-JUL-1997;   97US-0054217P.
PR   30-JUL-1997;   97US-0054218P.
PR   30-JUL-1997;   97US-0054234P.
PR   30-JUL-1997;   97US-0054236P.
PR   18-AUG-1997;   97US-0055968P.
PR   18-AUG-1997;   97US-0055969P.
PR   18-AUG-1997;   97US-0055972P.
PR   19-AUG-1997;   97US-0056534P.
PR   19-AUG-1997;   97US-0056543P.
PR   19-AUG-1997;   97US-0056554P.
PR   19-AUG-1997;   97US-0056561P.
PR   19-AUG-1997;   97US-0056727P.
PR   19-AUG-1997;   97US-0056729P.
PR   19-AUG-1997;   97US-0056730P.
PR   29-JUL-1998;   98WO-US015949.
PR   26-JAN-1999;   99US-00236557.
PR   21-SEP-2000; 2000US-00666987.
XX
PA   (RUBE/) RUBEN S M.
PA   (FENG/) FENG P.
PA   (LAFL/) LAFLEUR D W.
PA   (MOOR/) MOORE P A.
PA   (SHIY/) SHI Y.
PA   (KYAW/) KYAW H.
PA   (LIYY/) LI Y.
PA   (ZENG/) ZENG Z.
PA   (CART/) CARTER K C.
PA   (ENDR/) ENDRESS G A.
PA   (WEIY/) WEI Y.
PA   (FANP/) FAN P.
PA   (ROSE/) ROSEN C A.
XX
PI   Ruben SM,  Feng P,  Lafleur DW,  Moore PA,  Shi Y,  Kyaw H,  Li Y;
PI   Zeng Z,  Carter KC,  Endress GA,  Wei Y,  Fan P,  Rosen CA;
XX
DR   WPI; 2003-852813/79.
DR   P-PSDB; ADD90301.
XX
PT   New nucleic acid molecule, useful for preparing a medicament for
PT   preventing, treating or ameliorating a medical condition e.g., cancer.
XX
PS   Claim 1; SEQ ID NO 25; 213pp; English.
XX
CC   The invention describes novel isolated human nucleic acids. The nucleic
CC   acid is useful for preparing a medicament for preventing, treating or
```

```
CC   ameliorating a medical condition e.g., cancer, and in gene therapy. This
CC   sequence encodes a novel human secreted protein of the invention.
XX
SQ   Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;

  Query Match             31.4%;  Score 1002.4;  DB 10;  Length 1510;
  Best Local Similarity   99.2%;  Pred. No. 4.6e-167;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
              ||||||||||||||| |||
Db       1132 AAATGCCAAAAGGCATTT 1149


RESULT 5
ADG90029
ID   ADG90029 standard; cDNA; 1510 BP.
XX
AC   ADG90029;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human cDNA from secreted protein gene 15.
XX
KW   Secreted protein; gene therapy; neural disorder; immune system disorders;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; cancer; systemic lupus erythematosus;
KW   rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia;
KW   Grave's disease; diabetes; hepatitis; asthma; allergy; nephritis;
KW   Parkinson's disease; Alzheimer's disease; atherosclerosis;
KW   myocardial infarction; AIDS; infection; human; ss; gene.
XX
OS   Homo sapiens.
XX
PN   US2003166541-A1.
XX
PD   04-SEP-2003.
XX
PF   04-JUN-2002; 2002US-00160162.
XX
PR   30-JUL-1997;   97US-0054209P.
PR   30-JUL-1997;   97US-0054211P.
PR   30-JUL-1997;   97US-0054212P.
PR   30-JUL-1997;   97US-0054213P.
PR   30-JUL-1997;   97US-0054214P.
PR   30-JUL-1997;   97US-0054215P.
PR   30-JUL-1997;   97US-0054217P.
PR   30-JUL-1997;   97US-0054218P.
PR   30-JUL-1997;   97US-0054234P.
PR   30-JUL-1997;   97US-0054236P.
PR   18-AUG-1997;   97US-0055968P.
PR   18-AUG-1997;   97US-0055969P.
PR   18-AUG-1997;   97US-0055972P.
PR   19-AUG-1997;   97US-0056534P.
PR   19-AUG-1997;   97US-0056543P.
PR   19-AUG-1997;   97US-0056554P.
PR   19-AUG-1997;   97US-0056561P.
PR   19-AUG-1997;   97US-0056727P.
PR   19-AUG-1997;   97US-0056729P.
PR   19-AUG-1997;   97US-0056730P.
PR   29-JUL-1998;   98WO-US015949.
PR   26-JAN-1999;   99US-00236557.
PR   05-JUN-2001; 2001US-0295558P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Feng P,  Lafleur DW,  Moore PA,  Shi Y,  Kyaw H,  Li Y;
PI   Zeng Z,  Carter KC,  Endress GA,  Wei Y,  Fan P,  Rosen CA;
XX
DR   WPI; 2003-874923/81.
DR   P-PSDB; ADG90120.
XX
PT   Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
PT   diagnosing and treating disorders related to their aberrant expression
PT   and activity.
XX
PS   Claim 1; SEQ ID NO 25; 308pp; English.
```

```
XX
CC   The invention relates to an isolated nucleic acid molecule encoding a
CC   secreted protein that is at least 95% identical to a polynucleotide
CC   fragment of any of the nucleotide sequences listed in table 1A of the
CC   specification, which is hybridisable to the nucleotide sequences, a
CC   polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
CC   or epitope of any of the amino acid sequences) listed in table 1A of the
CC   specification, a polynucleotide which is an (allelic) variant of the
CC   nucleotide sequences listed in the specification, a polynucleotide which
CC   encodes a species homologue of the above amino acid sequences, a
CC   polynucleotide capable of hybridising under stringent conditions to any
CC   of the above polynucleotides, where the polynucleotide does not hybridise
CC   under stringent conditions to a nucleic acid molecule having a nucleotide
CC   sequence of only A or T residues. Also included are a recombinant vector
CC   comprising the above nucleic acid molecule, making a recombinant host
CC   cell comprising the above nucleic acid molecule, an isolated polypeptide
CC   comprising a sequence that is at least 95% identical to the polypeptide
CC   (or its fragment, domain, epitope, secreted form, (allelic) variant or
CC   homologue) encoded by the above nucleic acid molecule, an isolated
CC   antibody that binds specifically to the above polypeptide, a recombinant
CC   host cell produced by the above method and that expresses the above
CC   polypeptide, making an isolated polypeptide, preventing, treating or
CC   ameliorating a medical condition, diagnosing a pathological condition or
CC   a susceptibility to a pathological condition in a subject, identifying a
CC   binding partner to the above polypeptide, the gene corresponding to the
CC   cDNA sequence given in the specification, and identifying an activity in
CC   a biological assay. The nucleic acid molecule and polypeptide are useful
CC   in diagnosing, preventing, prognosing or treating diseases or disorders
CC   associated with aberrant expression and/or activity of the above
CC   polypeptide, such as neural disorders, immune system disorders, muscular
CC   disorders, reproductive disorders, gastrointestinal disorders, pulmonary
CC   disorders, cardiovascular disorders, renal disorders, proliferative
CC   disorders and/or cancers. In particular, these diseases are systemic
CC   lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
CC   thyroiditis, anaemia, Grave's disease, diabetes, hepatitis, asthma,
CC   allergies, nephritis, Parkinson's disease, Alzheimer's disease,
CC   atherosclerosis, myocardial infarction, AIDS and infections. The methods
CC   may be used for identifying agonists and antagonists of the
CC   polynucleotide and polypeptide. The present sequence is a cDNA from one
CC   of the 83 disclosed secreted protein genes.
XX
SQ   Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;

  Query Match             31.4%;  Score 1002.4;  DB 10;  Length 1510;
  Best Local Similarity   99.2%;  Pred. No. 4.6e-167;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             ||||||||||||||| |||
Db       1132 AAATGCCAAAAGGCATTT 1149

RESULT 6
ADQ24238/c
ID   ADQ24238 standard; DNA; 2058 BP.
XX
AC   ADQ24238;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 7058.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
```

```
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 7058; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 2058 BP; 488 A; 468 C; 405 G; 695 T; 0 U; 2 Other;

  Query Match             31.4%;  Score 1002.4;  DB 12;  Length 2058;
  Best Local Similarity   99.2%;  Pred. No. 5e-167;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1739

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1679

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1619

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1559

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 1499

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 1439

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 1379

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 1319

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 1259

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 1199

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 1139

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 1079

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 1019

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       1018 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 959

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 899

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 839

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        838 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 779

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db        778 AAATGCCAAAAGGCATTT 761


RESULT 7
ADJ75198/c
ID   ADJ75198 standard; DNA; 1321 BP.
XX
AC   ADJ75198;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Marker gene SEQ ID NO:450.
XX
KW   bronchial asthma; chronic obstructive pulmonary disease;
KW   respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW   gene therapy; marker gene; gene; ds.
XX
OS   Homo sapiens.
XX
PN   EP1394274-A2.
XX
PD   03-MAR-2004.
XX
PF   04-AUG-2003; 2003EP-00254857.
XX
PR   06-AUG-2002; 2002JP-00229312.
PR   20-MAR-2003; 2003JP-00077212.
XX
PA   (GENO-) GENOX RES INC.
XX
PI   Ohtani N,  Sugita Y,  Yamaya M,  Kubo H,  Nagai H,  Izuhara K;
XX
DR   WPI; 2004-193155/19.
XX
PT   Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT   comparing the expression level of a marker gene in a biological sample
PT   from a subject with the expression level of the gene in a sample from a
PT   healthy subject.
XX
PS   Claim 1; SEQ ID NO 450; 241pp; English.
XX
CC   The present invention describes a method of testing for bronchial asthma
CC   or chronic obstructive pulmonary disease. The method comprises
CC   determining the expression level of a marker gene in a biological sample
CC   from a subject, comparing the expression level determined with the
CC   expression level of the marker gene in a biological sample from a healthy
```

```
CC   subject, and judging whether the subject has bronchial asthma or chronic
CC   obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC   genes (S1) whose expression levels increase when respiratory epithelial
CC   cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC   whose expression levels decrease when respiratory epithelial cells are
CC   stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC   testing for bronchial asthma or chronic obstructive pulmonary disease;
CC   (2) a kit for screening for a candidate compound for a therapeutic agent
CC   to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC   an animal model for bronchial asthma or chronic obstructive pulmonary
CC   disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC   method for producing an animal model for bronchial asthma or chronic
CC   obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC   asthma or chronic obstructive pulmonary disease, comprising the compound,
CC   a marker gene or an antisense nucleic acid corresponding to a portion of
CC   the marker gene, a ribozyme, a polynucleotide that suppresses the
CC   expression of the gene through an RNAi effect or an antibody recognising
CC   a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC   bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC   probe has been immobilised to assay a marker gene. (I) has respiratory
CC   and antiasthmatic activities, and can be used in gene therapy. The method
CC   is useful for testing for or screening for a therapeutic agent for
CC   bronchial asthma or chronic obstructive pulmonary disease. The present
CC   sequence is used in the exemplification of the present invention.
XX
SQ   Sequence 1321 BP; 414 A; 255 C; 236 G; 416 T; 0 U; 0 Other;

  Query Match             31.4%;  Score 1000.8;  DB 12;  Length 1321;
  Best Local Similarity   99.1%;  Pred. No. 8.6e-167;
  Matches 1029;  Conservative    0;  Mismatches    2;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1230

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1170

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1110

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1050

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 990

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 930

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 870

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 810

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 750

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 690

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        689 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 630

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 570

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        569 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 510

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        509 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 450

Qy       2992 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 ACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 390

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 330

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        329 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 270

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db        269 AAATGCCAAAAGGCATTT 252
```

```
RESULT 8
ADN04817/c
ID   ADN04817 standard; cDNA; 1321 BP.
XX
AC   ADN04817;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic cDNA sequence #623.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   P-PSDB; ADN04818.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 1; SEQ ID NO 1211; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polynucleotides of the invention.
XX
```

```
SQ   Sequence 1321 BP; 414 A; 255 C; 236 G; 416 T; 0 U; 0 Other;

  Query Match             31.4%;  Score 1000.8;  DB 12;  Length 1321;
  Best Local Similarity   99.1%;  Pred. No. 8.6e-167;
  Matches 1029;  Conservative    0;  Mismatches    2;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1230

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1170

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1110

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1050

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 990

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 930

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 870

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 810

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 750

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 690

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 630

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 570

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 510

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        509 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 450

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 390

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 330

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db        329 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 270

Qy       3171 AAATGCCAAAAGGCTTTT 3188
              |||||||||||||| |||
Db        269 AAATGCCAAAAGGCATTT 252


RESULT 9
ADR25935
ID   ADR25935 standard; DNA; 1321 BP.
XX
AC   ADR25935;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Breast cancer prognosis marker #1796.
XX
KW   ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004065545-A2.
XX
PD   05-AUG-2004.
XX
PF   15-JAN-2004; 2004WO-US001100.
XX
PR   15-JAN-2003; 2003US-00342887.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
PA   (NECA-) NETHERLANDS CANCER INST.
XX
PI   Van't Veer LJ,  He Y;
XX
DR   WPI; 2004-593473/57.
XX
PT   Classifying a breast cancer patient according to prognosis comprises
PT   determining the similarity between the level of expression of each of
PT   five genes in a cell sample taken from patient, to control levels.
XX
PS   Disclosure; SEQ ID NO 1796; 226pp; English.
XX
CC   The invention relates to a method of classifying a breast cancer patient
CC   according to prognosis by determining the similarity between the level of
CC   expression of each of five genes for which markers are listed in the
CC   specification, in a cell sample taken from the breast cancer patient, to
CC   control levels of expression for each respective five genes to obtain a
CC   patient similarity value. The methods are useful for classifying a breast
CC   cancer patient according to prognosis. Kits and computer program products
CC   are useful for data analysis using the diagnostic, prognostic and
CC   statistical methods of the invention. This sequence corresponds to a
CC   marker used in the method of the invention.
XX
SQ   Sequence 1321 BP; 416 A; 236 C; 255 G; 414 T; 0 U; 0 Other;

  Query Match             31.4%;  Score 1000.8;  DB 13;  Length 1321;
  Best Local Similarity   99.1%;  Pred. No. 8.6e-167;
  Matches 1029;  Conservative    0;  Mismatches    2;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          33 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 92

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          93 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 152

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         153 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 212

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         213 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 272
```

```
Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 332

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 392

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 452

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 512

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 572

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 632

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 692

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 752

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        753 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 812

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        813 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 872

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 932

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 992

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        993 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1052

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db       1053 AAATGCCAAAAGGCATTT 1070
```

```
RESULT 10
AAC87358
ID   AAC87358 standard; cDNA; 1581 BP.
XX
AC   AAC87358;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human developmentally-regulated hippocampus EST, SEQ ID NO:1.
XX
KW   Human hippocampal gene; hippocampus EST; expressed sequence tag;
KW   developmental regulation; drug screening; modulator;
KW   expression alteration; Alzheimer's disease; Huntington's disease;
KW   schizophrenia; epilepsy; diagnostic probe; gene therapy; nootropic;
KW   neuroprotective; anticonvulsant; neuroleptic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200070036-A2.
XX
PD   23-NOV-2000.
XX
PF   10-MAY-2000; 2000WO-US013046.
XX
PR   17-MAY-1999;   99US-00313300.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Kaser MR,  Lal P,  Yue H,  Tang YT,  Baughn MR,  Azimzai Y;
XX
DR   WPI; 2001-016229/02.
XX
PT   Nucleic acids expressed in hippocampus useful for diagnosing, treating or
PT   preventing diseases associated with the hippocampus, e.g., Alzheimer's
PT   disease, Huntington's disease, Schizophrenia and epilepsy.
XX
PS   Claim 2; Page 31; 39pp; English.
XX
CC   The invention relates to novel human developmentally-regulated genes
CC   which are expressed in the hippocampus (ESTs given in AAC87358-C87364),
CC   and to a protein product of a gene of the invention (AAB29800). The human
CC   developmentally-regulated hippocampal gene ESTs (expressed sequence tags)
CC   were identified on the basis of homology with ESTs (AAC87365-C87371)
CC   found in adult rat hippocampus but not in foetal rat hippocampus. The
CC   invention also relates to expression vectors and host cells comprising a
CC   human developmentally-regulated hippocampal gene; recombinant production
CC   of the gene product; and a method of screening for a compound which
CC   specifically binds to a human hippocampal gene of the invention or its
CC   product. It additionally encompasses nucleic acid sequences with at least
CC   70% identity to one of the human hippocampal ESTs, and proteins with at
CC   least 85% identity to the human hippocampal protein of the invention. The
CC   human hippocampal genes, their encoded proteins and modulators of the
CC   genes or proteins are useful for treating or preventing diseases
CC   associated with altered expression of a gene in the hippocampus, such as
CC   Alzheimer's disease, Huntington's disease, schizophrenia, epilepsy and
CC   complications of these conditions. The nucleic acid sequences, or
CC   fragments thereof are also as hybridisation probes for use in diagnosing
CC   a disease or condition associated with altered expression of a
CC   hippocampal gene. The present sequence represents a human hippocampal EST
CC   corresponding to a developmentally-regulated gene of the invention
XX
SQ   Sequence 1581 BP; 500 A; 307 C; 325 G; 449 T; 0 U; 0 Other;
```

```
  Query Match             31.3%;  Score 998.4;  DB 4;  Length 1581;
  Best Local Similarity   98.8%;  Pred. No. 2.4e-166;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels   11;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    :   204 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 263

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 323

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 383

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 443

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 503

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 563

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 623

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 683

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 743

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 803

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 863

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 923

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 983

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA----------AC 2987
                ||||||||||||||||||||||||||||||||||||||||||||||||          ||
Db        984 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1043

Qy       2988 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 3047
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 1103

Qy       3048 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 3107
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1104 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 1163

Qy       3108 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-T 3166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1164 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATAT 1223

Qy       3167 CTTGAAATGCCAAAAGGCTTTT 3188
                ||||||||||||||||| |||
Db       1224 CTTGAAATGCCAAAAGGCATTT 1245
```

RESULT 11
ADI61702
ID ADI61702 standard; cDNA; 1876 BP.
XX
AC ADI61702;
XX
DT 22-APR-2004 (first entry)
XX
DE Human cDNA downregulated in Alzheimer's disease, INCYTE 234151.1.
XX
KW Human; ss; Alzheimer's disease; differential display; neuroprotective;
KW brain disorder.
XX
OS Homo sapiens.
XX
PN US6682888-B1.
XX
PD 27-JAN-2004.
XX
PF 05-MAY-2000; 2000US-00566921.
XX
PR 05-MAY-2000; 2000US-00566921.
XX
PA (INCY-) INCYTE CORP.
XX
PI Loring JF, Tingley DW, Edwards CM;
XX
DR WPI; 2004-118572/12.
XX
PT New composition comprising cDNAs that are differentially expressed in
PT brain disorders, useful for diagnosing or treating Alzheimer's disease.
XX
PS Claim 1; SEQ ID NO 70; 223pp; English.
XX
CC The invention relates to a new composition comprising ADI61633-
CC ADI61770and their complements that are cDNAs differentially expressed in
CC brain disorders. Also included are a high throughput method for detecting
CC differential expression of one or more cDNAs in a sample containing
CC nucleic acids and a high throughput method for screening a library of
CC molecules or compounds to identify a ligand that specifically binds a
CC cDNA. The expression of the each of the cDNAs is downregulated at least
CC two-fold in the brain of the subjects with Alzheimer's disease (ADI61633-
CC ADI61727) or upregulated at least two fold in Alzheimer's disease
CC (ADI61728-ADI61770). The composition is useful for diagnosing or treating
CC Alzheimer's disease. The present sequence is a cDNA downregulated at
CC least two-fold in the brain of the subjects with Alzheimer's disease.
XX
SQ Sequence 1876 BP; 480 A; 442 C; 442 G; 480 T; 0 U; 32 Other;

Query Match 30.2%; Score 963.8; DB 12; Length 1876;
Best Local Similarity 96.5%; Pred. No. 3.1e-160;
Matches 1004; Conservative 0; Mismatches 27; Indels 9; Gaps 3;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 579

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 639

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 699

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 759

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAAT-AAAAAAATAA 2456
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 AGGAGGGGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAAATAA 819

Qy       2457 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 2516
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 879

Qy       2517 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 2576
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 939

Qy       2577 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 2636
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 999

Qy       2637 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 2696
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 1059

Qy       2697 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTAT 2756
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTAT 1119

Qy       2757 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 2816
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 1179

Qy       2817 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 2876
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 1239

Qy       2877 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 2936
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 1299

Qy       2937 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 2996
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 1359

Qy       2997 -------CACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCT 3049
                                                    ||||||||||||||||||||||||||||
Db       1360 NGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGAAAAACAAAGAAAAAAATGCTTGAGCT 1419

Qy       3050 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 3109
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 1479

Qy       3110 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCT 3168
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1480 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCT 1539

Qy       3169 TGAAATGCCAAAAGGCTTTT 3188
                ||||||||||||||||| |||
Db       1540 TGAAATGCCAAAAGGCATTT 1559
```

RESULT 12
ABQ60975
ID ABQ60975 standard; cDNA; 873 BP.
XX
AC ABQ60975;
XX
DT 26-FEB-2003 (first entry)
XX
DE PRO1489 protein encoding sequence.
XX
KW Neuroprotective; immunomodulator; cancer; chromosome 1; cytostatic;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnerary; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231111-A2.
XX
PD 18-APR-2002.
XX
PF 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
DR N-PSDB; ABP43731.
XX
PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
PS Claim 1; SEQ ID # 188; 357pp + Sequence Listing; English.
XX

CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerary, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABQ60788-
CC ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 873 BP; 304 A; 144 C; 174 G; 251 T; 0 U; 0 Other;

  Query Match   24.8%; Score 791.4; DB 6; Length 873;
  Best Local Similarity 99.6%; Pred. No. 6.3e-130;
  Matches 814; Conservative   0; Mismatches   1; Indels   2; Gaps   2;

```
Qy       2205 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 2264
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 112

Qy       2265 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 2324
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     :   113 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 172

Qy       2325 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 2384
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       173 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 232

Qy       2385 ATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAA 2444
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       233 ATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAA 292

Qy       2445 ATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAA 2504
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       293 ATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAA 352

Qy       2505 ATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATA 2564
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       353 ATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATA 412

Qy       2565 CGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACC 2624
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       413 CGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACC 472

Qy       2625 AGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATT 2684
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       473 AGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATT 532

Qy       2685 TAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTT 2744
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       533 TAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTT 592

Qy       2745 TTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGT 2804
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       593 TTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGT 652

Qy       2805 TTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTT 2864
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       653 TTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGACGAGAACTTTT 712

Qy     : 2865 CTAAGGAATG-ATACATAG-AAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTT 2922
                |||||||||  |||||||| ||||||||||||||||||||||||||||||||||||||||
Db       713 CTAAGGAATGCATACATAGCAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTT 772

Qy       2923 CTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGA 2982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       773 CTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGA 832
```

```
Qy       2983 AAAACACACACACACACACACACACACACACACACAC 3019
              ||||||||||||||||||||||||||||||||||||||
Db        833 AAAACACACACACACACACACACACACACACACACAC 869


RESULT 13
AAD31052
ID   AAD31052 standard; cDNA; 1104 BP.
XX
AC   AAD31052;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human PAPAP cDNA.
XX
KW   Human; PAPAP protein; schizophrenia candidate gene; g34872 gene;
KW   schizophrenia; bipolar disorder; central nervous system disorder;
KW   psychotic disorder; mood disorder; autism; mental retardation;
KW   psychiatric disease; anxiety disorder; impulse-control disorder;
KW   eating disorder; cognitive disorder; personality disorder; vaccine;
KW   chromosome 1p35-p36; neuroleptic; antialcoholic; tranquiliser;
KW   antidepressant; nootropic; antiaddictive; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .86
FT                   /*tag=  a
FT   CDS             87. .344
FT                   /*tag=  b
FT                   /product= "Human PAPAP protein"
FT   3'UTR           345. .1104
FT                   /*tag=  c
FT   polyA_site      1085. .1104
FT                   /*tag=  d
XX
PN   WO200212279-A2.
XX
PD   14-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-IB001891.
XX
PR   07-AUG-2000; 2000US-0223482P.
XX
PA   (GEST ) GENSET.
XX
PI   Bihain B,  Bour B,  Bougueleret L;
XX
DR   WPI; 2002-241732/29.
DR   P-PSDB; AAE19431.
XX
PT   Novel isolated and purified or recombinant polynucleotide encoding PAPAP
PT   protein, useful for diagnosing and treating schizophrenia, bipolar
PT   disorder and other central nervous system disorders.
XX
PS   Claim 2; Page 92; 96pp; English.
XX
CC   The invention relates to human PAPAP polypeptides and polynucleotides.
CC   The invention also concerns the interaction of PAPAP with schizophrenia
CC   candidate gene g34872. PAPAP polypeptides, gene and anti-PAPAP antibodies
CC   are useful for treating schizophrenia, bipolar disorder or related
CC   central nervous system (CNS) disorders e.g. psychotic disorders, mood
CC   disorders, autism, substance dependence and alcoholism, mental
CC   retardation and other psychiatric diseases including cognitive, anxiety,
CC   eating, impulse-control and personality disorders. They are also used in
CC   vaccines. The present sequence is human PAPAP cDNA. The PAPAP gene is
CC   located on chromosome 1p35-p36
XX
SQ   Sequence 1104 BP; 350 A; 220 C; 262 G; 272 T; 0 U; 0 Other;

  Query Match           19.0%;  Score 607.2;  DB 6;  Length 1104;
  Best Local Similarity  89.2%;  Pred. No. 2e-97;
  Matches  787;  Conservative    0;  Mismatches   63;  Indels   32;  Gaps   11;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 311

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              |||||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db        312 -CTGGTGTCTAACT-CCCCAAAGACAATGAGTTAAGGGAGAGAATAGGAACGGCGGTAAC 369

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        370 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TAC 428

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              |||||||||||||||||||||||||||||| | ||| |||||||||| |||||| |||
Db        429 CCACGATGAAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGGCGA-AGG 486

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              |||| |||   ||| |||| |  ||||| ||||| |||| | || || ||||||| |
Db        487 AGGAAGAGAGAGAGAAGAAAGGCTTGGGC-CCTCTACAAATAAAATAAAAAAAAAAAATT 545

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
              ||||||  |||||    || |                  | | ||  | |||  ||||
Db        546 TAAAATAATAAAATCCCTATA----------------TCCCATATAAGAATAAAAGAGTC 589

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
              |||||    ||  | |    ||| | | ||||||||||||||    ||     ||
Db        590 TCAGT----GCAGTATTGGCAAAATTAAATCCATTTCTTTTTAATACGGG----AATATT 641

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
              |  |||||||||||||||||| |||||||||||||||| || ||||||||||||||||
Db        642 GGCATTATAGATCTGGATTTTG-ACCACTTAATGAAGCGGC-ACCCCAGGTGTTTTGAGG 699

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 759

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 819

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 879

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 939

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 999

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 2997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1059

Qy       2998 ACACACACACACACACACACACACGAAAAACAAAGAAAAAAA 3039
              |||||||||||||||||||||||||||| ||| |||||||
Db       1060 ACACACACACACACACACACACACGAAAAAAAAAAAAAAAA 1101


RESULT 14
ABQ43206/c
ID   ABQ43206 standard; DNA; 869 BP.
XX
AC   ABQ43206;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 29797.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
```

KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 869 BP; 106 A; 123 C; 361 G; 279 T; 0 U; 0 Other;

Query Match 18.3%; Score 583.8; DB 6; Length 869;
Best Local Similarity 79.6%; Pred. No. 2.5e-93;
Matches 690; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

```
Qy        51 TCTGCCTCCCCGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTC 110
             ||| |||||||||||||| ||||||||||||||||||||||||  |||| ||||||||| |||
Db       868 TCTACCTCCCCGAACCGACGCCGCCGCCCGCACTCGCCGCAAAACCGACCCGCCCGACTC 809

Qy       111 CCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCC 170
             |||   | ||||||||||| ||||||||||||||  ||||||   ||||||||||||||||
Db       808 CCGAAATACGCCCTCCTCGATCCCGCGCCCTCCGAACTCGCAAAAACGTCTCCTCCCTCC 749

Qy       171 CGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGA 230
             || |||||| ||||||||| |||| |||||||||| | ||||| ||||||||  || | |
Db       748 CGACTCGCGACCCCGCCCGACCCGACCCCCGCCCAAAACCCCAACGCGCCGAAAATATAA 689

Qy       231 GTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAG 290
             |||| |||||||||  || | || || ||||||||||||| | ||| | |||| ||||
Db       688 ATCCTACTCGCCTCTAACGAAACAACAACCACTCGCGCGCGAAACCGAAACGCAACGCAA 629

Qy       291 CGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCG 350
             |||| |||||  |||||||||  |||||||||||| | |||||||||| ||||| || ||
Db       628 CGCAACCGCGAACGCTCTCCGAACCGCTCGCGCGAATACCGCGCTCTTACCCTAACGACG 569

Qy       351 TCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCC 410
             ||||||| |||||||||| |||||||||||||| || ||||||  |||||||| ||| ||
Db       568 TCCCCCGACCTCTCGCCGACGCCACCGCCGCAACAACCCGCGAACCGTCCCCGACCGACC 509

Qy       411 GCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGG 470
             ||||||| ||||| ||||||| |||| ||||||||  ||   ||||| ||| | |
Db       508 GCCCCCGACCCCAACGCCGCTAACCCTATCCGCCGCGAACGAAAACGCGATCGAAAAAAA 449

Qy       471 CGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAA 530
             ||||||| || | |||||| |||||||||| ||| |  | || ||||||| |||||| ||
Db       448 CGCCGCGACGAAACCCCCGAACGCGACCATATCGAAAATACTACCCTACGACGACGAAAA 389

Qy       531 GCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGA 590
             || | |||||||| ||||| || ||||||   ||| |||||||||| |||||| ||  |
Db       388 ACTAAACCCCTACGACGACGACGACGACGTAAACCAAATCTTCTCCTACCGCCTACAAAA 329

Qy       591 CACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGG 650
             |||||||||||||||| |||||  || ||||| || |||||||| ||   ||| ||||
Db       328 CACCAACAACTTCTTCGACGCCGAACAAAACAAACGACCGCCCAAACTAAACCAAATCGA 269

Qy       651 CCGGAGCAAGCGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCG 710
             ||| | ||| ||   | | |||||| ||||||| ||| ||||||||||||||||||||||
Db       268 CCGAAACAAACGAAATAAATTCGCGACCCCCTTATCTAACACCCCCTTTTTCCCGCGCCG 209

Qy       711 CGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGA 770
             || ||| |||||   || || |  ||||||||||||||  | |||||||| |||| ||| |
Db       208 CGACCTAAACAAAAATTACGAAAATCTCCCACCCGCTAAAACCCGTTCAAACCTAACGAA 149

Qy       771 ATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGG 830
             |||||||||| || ||||     |||||||||| ||  |||| || || ||  |||||
Db       148 ATCCCTTCTTACAAAATTAAAAAATCCCGCACTACGAATCCGACTAAAACGAATCGCAAA 89

Qy       831 AACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAG 890
             |||||||||||||||| ||| ||||||| ||   |||||||     |||  || ||||||
Db        88 AACGCGTCCCCCTAAACCGAATCCCCGACTAAATCACCCTAAAAACGTAACGACTTCTAA 29

Qy       891 CAGCAGCTGGGGGTCTCCACCCGCGCG 917
             || || ||      ||||||||||||||
Db        28 CAACAACTAAAAATCTCCACCCGCGCG 2
```

RESULT 15
ABQ43207
ID ABQ43207 standard; DNA; 869 BP.
XX
AC ABQ43207;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29798.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.

XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 869 BP; 279 A; 361 C; 123 G; 106 T; 0 U; 0 Other;

Query Match 18.3%; Score 583.8; DB 6; Length 869;
Best Local Similarity 79.6%; Pred. No. 2.5e-93;
Matches 690; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

```
Qy         51 TCTGCCTCCCCGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTC 110
              ||| |||||||||||||| |||||||||||||||||||||||||  |||| |||||||| |||
Db          2 TCTACCTCCCCGAACCGACGCCGCCGCCCGCACTCGCCGCAAAACCGACCCGCCCGACTC 61

Qy        111 CCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCC 170
              |||   | |||||||||||| |||||||||||||  ||||||  ||||||||||||||||
Db         62 CCGAAATACGCCCTCCTCGATCCCGCGCCCTCCGAACTCGCAAAAACGTCTCCTCCCTCC 121

Qy        171 CGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGA 230
              || |||||| |||||||||| |||| |||||||||| | ||||| ||||||||  || | |
Db        122 CGACTCGCGACCCCGCCCGACCCGACCCCCGCCCAAAACCCCAACGCGCCGAAAATATAA 181

Qy        231 GTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAG 290
               |||| |||||||||  || | || || |||||||||||||| | ||| | |||| ||||
Db        182 ATCCTACTCGCCTCTAACGAAACAACAACCACTCGCGCGCGAAACCGAAACGCAACGCAA 241

Qy        291 CGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCG 350
              |||| |||||  |||||||||  |||||||||||| | |||||||||| ||||| || ||
Db        242 CGCAACCGCGAACGCTCTCCGAACCGCTCGCGCGAATACCGCGCTCTTACCCTAACGACG 301

Qy        351 TCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCC 410
              ||||||| |||||||||| |||||||||||||| || ||||||  ||||||||| ||| ||
Db        302 TCCCCCGACCTCTCGCCGACGCCACCGCCGCAACAACCCGCGAACCGTCCCCGACCGACC 361

Qy        411 GCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGG 470
              ||||||| ||||| ||||||| ||||| |||||||||  ||   ||||| ||| |  |
Db        362 GCCCCCGACCCCAACGCCGCTAACCCTATCCGCCGCGAACGAAAACGCGATCGAAAAAAA 421

Qy        471 CGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAA 530
              ||||||| || | |||||| |||||||||| ||| |  | || ||||||| |||||| ||
Db        422 CGCCGCGACGAAACCCCCGAACGCGACCATATCGAAAATACTACCCTACGACGACGAAAA 481

Qy        531 GCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGA 590
              || | |||||||| ||||| || ||||||   ||| |||||||||| |||||| ||  |
Db        482 ACTAAACCCCTACGACGACGACGACGACGTAAACCAAATCTTCTCCTACCGCCTACAAAA 541

Qy        591 CACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGG 650
              |||||||||||||||| |||||  || ||||| || |||||||| ||   ||| ||||
Db        542 CACCAACAACTTCTTCGACGCCGAACAAAACAAACGACCGCCCAAACTAAACCAAATCGA 601

Qy        651 CCGGAGCAAGCGGGGTGAGTTCGCGGCCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCG 710
              ||| | ||| ||   | | |||||| ||||||| ||| ||||||||||||||||||||||
Db        602 CCGAAACAAACGAAATAAATTCGCGACCCCCTTATCTAACACCCCCTTTTTCCCGCGCCG 661

Qy        711 CGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGA 770
              || ||| |||||   || || |  |||||||||||||  | |||||||| |||| ||| |
Db        662 CGACCTAAACAAAAATTACGAAAATCTCCCACCCGCTAAAACCCGTTCAAACCTAACGAA 721

Qy        771 ATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGG 830
              |||||||||| || ||||     |||||||||| ||  |||| || || ||  |||||
Db        722 ATCCCTTCTTACAAAATTAAAAAATCCCGCACTACGAATCCGACTAAAACGAATCGCAAA 781

Qy        831 AACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAG 890
              |||||||||||||||| ||| ||||||| ||   |||||||     |||  || ||||||
Db        782 AACGCGTCCCCCTAAACCGAATCCCCGACTAAATCACCCTAAAAACGTAACGACTTCTAA 841

Qy        891 CAGCAGCTGGGGGTCTCCACCCGCGCG 917
              || || ||     |||||||||||||
Db        842 CAACAACTAAAAATCTCCACCCGCGCG 868
```

Search completed: March 27, 2005, 22:02:23
Job time : 1548.62 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 27, 2005, 21:07:42 ; Search time 531.129 Seconds
(without alignments)
9824.526 Million cell updates/sec

Title: US-10-071-645-3
Perfect score: 3189
Sequence: 1 ccctcccctcccccctccgcc.....gaaatgccaaaaggctttta 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	998.4	31.3	1581	3	US-09-313-300-1	Sequence 1, Appli
	2	963.8	30.2	1876	4	US-09-566-921-70	Sequence 70, Appl
	3	507	15.9	524	4	US-09-513-999C-10663	Sequence 10663, A
c	4	413.4	13.0	417	4	US-09-702-705-475	Sequence 475, App
c	5	413.4	13.0	417	4	US-09-736-457-475	Sequence 475, App
c	6	413.4	13.0	417	4	US-09-614-124B-475	Sequence 475, App
c	7	413.4	13.0	417	4	US-09-671-325-475	Sequence 475, App
c	8	413.4	13.0	417	4	US-09-589-184-475	Sequence 475, App
c	9	413.4	13.0	417	4	US-09-658-824-475	Sequence 475, App
	10	319	10.0	336	4	US-09-854-133-560	Sequence 560, App
	11	199	6.2	1045	3	US-09-313-300-9	Sequence 9, Appli
	12	157.6	4.9	407	4	US-09-513-999C-8212	Sequence 8212, Ap
	13	154.2	4.8	315	4	US-09-621-976-15158	Sequence 15158, A
	14	114	3.6	183	4	US-09-513-999C-10536	Sequence 10536, A
c	15	101.2	3.2	7218	1	US-08-232-463-14	Sequence 14, Appl
	16	95	3.0	2561	4	US-09-616-289-48	Sequence 48, Appl
	17	94	2.9	94	4	US-09-513-999C-29981	Sequence 29981, A
	18	90.4	2.8	1614	4	US-09-616-289-45	Sequence 45, Appl
	19	90.4	2.8	12425	4	US-09-616-289-50	Sequence 50, Appl
	20	85.8	2.7	152331	3	US-09-128-155-16	Sequence 16, Appl
	21	80.6	2.5	12001	1	US-08-458-568A-11	Sequence 11, Appl
	22	80.4	2.5	3957	4	US-10-237-551-193	Sequence 193, App
	23	80.4	2.5	154746	4	US-09-827-688-8	Sequence 8, Appli
c	24	80.4	2.5	154746	4	US-09-827-688-8	Sequence 8, Appli
	25	77.6	2.4	90	4	US-09-513-999C-24816	Sequence 24816, A
	26	76.2	2.4	114793	4	US-10-148-806-3	Sequence 3, Appli
c	27	75.8	2.4	97423	4	US-09-949-016-12742	Sequence 12742, A
c	28	75.8	2.4	97424	4	US-09-949-016-15576	Sequence 15576, A
c	29	74.6	2.3	670689	4	US-09-949-016-12505	Sequence 12505, A
c	30	74.6	2.3	670690	4	US-09-949-016-14207	Sequence 14207, A
	31	74	2.3	18662	4	US-09-949-016-14655	Sequence 14655, A
c	32	73.4	2.3	14340	4	US-09-949-016-16972	Sequence 16972, A
	33	73.2	2.3	9369	4	US-10-237-551-190	Sequence 190, App
	34	73.2	2.3	9369	4	US-10-237-551-247	Sequence 247, App
c	35	73	2.3	15695	4	US-09-949-016-15644	Sequence 15644, A
	36	72.2	2.3	11024	4	US-09-949-016-12120	Sequence 12120, A
	37	72.2	2.3	11024	4	US-09-949-016-14656	Sequence 14656, A
c	38	70.6	2.2	4488	3	US-08-406-030A-3	Sequence 3, Appli
c	39	70.4	2.2	8438	1	US-07-945-283-1	Sequence 1, Appli
c	40	70.2	2.2	320	3	US-09-165-264-7	Sequence 7, Appli
c	41	70.2	2.2	1980	4	US-09-902-540-9228	Sequence 9228, Ap
c	42	70.2	2.2	13234	4	US-09-902-540-986	Sequence 986, App
	43	70.2	2.2	42053	4	US-09-949-016-15924	Sequence 15924, A
c	44	69.6	2.2	318	3	US-09-165-264-12	Sequence 12, Appl
c	45	69.6	2.2	319	3	US-09-165-264-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-313-300-1
; Sequence 1, Application US/09313300
; Patent No. 6222027
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew, R.
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Tom, Y.
; APPLICANT: Baughn, Mariah, R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: MOLECULES EXPRESSED IN HIPPOCAMPUS
; FILE REFERENCE: PB-0012 US
; CURRENT APPLICATION NUMBER: US/09/313,300
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: 239240
; PUBLICATION INFORMATION:
US-09-313-300-1

Query Match 31.3%; Score 998.4; DB 3; Length 1581;
Best Local Similarity 98.8%; Pred. No. 4.5e-198;
Matches 1030; Conservative 0; Mismatches 1; Indels 11; Gaps 2;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 263

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 323

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 383

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 443

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 503
```

```
Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 563

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 623

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 683

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 743

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 803

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 863

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 923

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 983

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA----------AC 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||          ||
Db        984 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1043

Qy       2988 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 3047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 ACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAG 1103

Qy       3048 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 3107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1104 CTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATT 1163

Qy       3108 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-T 3166
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1164 ATGTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATAT 1223

Qy       3167 CTTGAAATGCCAAAAGGCTTTT 3188
             ||||||||||||||||| |||
Db       1224 CTTGAAATGCCAAAAGGCATTT 1245
```

```
RESULT 2
US-09-566-921-70
; Sequence 70, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
;  APPLICANT:  Loring, Jeanne F.
;  APPLICANT:  Tingley, Debora W.
;  APPLICANT:  Edwards, Carla M.
;  TITLE OF INVENTION:  GENES EXPRESSED IN ALZHEIMER'S DISEASE
;  FILE REFERENCE:  PA-0024 US
;  CURRENT APPLICATION NUMBER: US/09/566,921
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  138
;  SOFTWARE:  PERL Program
; SEQ ID NO 70
;   LENGTH: 1876
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6682888 234151.1
;   NAME/KEY: unsure
;   LOCATION: 1360, 1362-1391, 1864
;   OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-70

  Query Match            30.2%;  Score 963.8;  DB 4;  Length 1876;
  Best Local Similarity   96.5%;  Pred. No. 7.7e-191;
  Matches 1004;  Conservative    0;  Mismatches   27;  Indels    9;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 579

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 639

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 699

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 759

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAAT-AAAAAAATAA 2456
             |||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        760 AGGAGGGGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAAATAA 819

Qy       2457 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 2516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGT 879

Qy       2517 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 2576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTG 939

Qy       2577 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 2636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAG 999

Qy       2637 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 2696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAA 1059

Qy       2697 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTAT 2756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTAT 1119

Qy       2757 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 2816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 CCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTC 1179

Qy       2817 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 2876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 ATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGAT 1239

Qy       2877 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 ACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAA 1299

Qy       2937 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 2996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACA 1359

Qy       2997 -------CACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCT 3049
                      ||||||||||||||||||||||||||||||
Db       1360 NGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGAAAAACAAAGAAAAAAATGCTTGAGCT 1419

Qy       3050 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1420 TTTTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 1479

Qy       3110 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCT 3168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1480 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCT 1539

Qy       3169 TGAAATGCCAAAAGGCTTTT 3188
             |||||||||||||||| |||
Db       1540 TGAAATGCCAAAAGGCATTT 1559


RESULT 3
US-09-513-999C-10663
; Sequence 10663, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 10663
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 465
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 471
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 475
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 488
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 491
;   OTHER INFORMATION: m=a or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 497
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 510
;   OTHER INFORMATION: k=g or t
US-09-513-999C-10663

  Query Match             15.9%;  Score 507;  DB 4;  Length 524;
  Best Local Similarity   98.5%;  Pred. No. 5.4e-96;
  Matches  516;  Conservative    5;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       2558 TTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGG 2617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGG 60

Qy       2618 CAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTA 2677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTA 120

Qy       2678 CATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTT 2737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTT 180

Qy       2738 ATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGAT 2797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGAT 240

Qy       2798 ATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAG 2857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAG 300

Qy       2858 AACTTTTCTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTG 2917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTTTTCTAAGGAATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTG 360

Qy       2918 TGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATG 2977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATG 420

Qy       2978 TCAGAAAAACACACACACACACACACACACACACACACACACA--CGAAAAACAAAGAAA 3035
             ||||||||||||||||||||||||||||||||||||||||||  ||||| |||:|||||
Db        421 TCAGAAAAACACACACACACACACACACACACACACACACACRCGAAANACARAGAAA 480

Qy       3036 AAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTG 3079
             |||||||:||:|||||:||||||||||||:|||||||||||||||
Db        481 AAAATGCKTGMGCTTTYTCTAACTTCCCCKTGCAGTCTGTTGTG 524


RESULT 4
US-09-702-705-475/c
; Sequence 475, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C14
;  CURRENT APPLICATION NUMBER: US/09/702,705
;  CURRENT FILING DATE:  2000-10-30
;  NUMBER OF SEQ ID NOS:  1833
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(417)
;   OTHER INFORMATION: n = A,T,C or G
US-09-702-705-475

  Query Match             13.0%;  Score 413.4;  DB 4;  Length 417;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-76;
  Matches  414;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             |||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298
```

```
Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1
```

RESULT 5
US-09-736-457-475/c
; Sequence 475, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-475

Query Match 13.0%; Score 413.4; DB 4; Length 417;
Best Local Similarity 99.3%; Pred. No. 1.5e-76;
Matches 414; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             ||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298

Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1
```

RESULT 6
US-09-614-124B-475/c
; Sequence 475, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-614-124B-475

Query Match 13.0%; Score 413.4; DB 4; Length 417;
Best Local Similarity 99.3%; Pred. No. 1.5e-76;
Matches 414; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             ||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298

Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1


RESULT 7
US-09-671-325-475/c
; Sequence 475, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(417)
;   OTHER INFORMATION: n = A,T,C or G
US-09-671-325-475

  Query Match             13.0%;  Score 413.4;  DB 4;  Length 417;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-76;
  Matches  414;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             ||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298

Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1


RESULT 8
US-09-589-184-475/c
; Sequence 475, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C8
;  CURRENT APPLICATION NUMBER: US/09/589,184
;  CURRENT FILING DATE:  2000-06-05
;  NUMBER OF SEQ ID NOS:  827
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(417)
;   OTHER INFORMATION: n = A,T,C or G
US-09-589-184-475

  Query Match             13.0%;  Score 413.4;  DB 4;  Length 417;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-76;
  Matches  414;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             ||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298

Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1


RESULT 9
US-09-658-824-475/c
; Sequence 475, Application US/09658824
; Patent No. 6746846
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
```

```
;  FILE REFERENCE:  210121.478C11
;  CURRENT APPLICATION NUMBER: US/09/658,824
;  CURRENT FILING DATE:  2000-09-08
;  NUMBER OF SEQ ID NOS:  1788
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(417)
;   OTHER INFORMATION: n = A,T,C or G
US-09-658-824-475

  Query Match             13.0%;  Score 413.4;  DB 4;  Length 417;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-76;
  Matches  414;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAA 2505
             |||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        417 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATNCNCATATAAAGAAA 358

Qy       2506 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 2565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATAC 298

Qy       2566 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 2625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCA 238

Qy       2626 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 2685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTT 178

Qy       2686 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTT 2745
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        177 AATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGGAATGCTGTCCAGTTTTATTTTTTT 118

Qy       2746 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 2805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTT 58

Qy       2806 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          57 TTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTT 1


RESULT 10
US-09-854-133-560
; Sequence 560, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 560
;   LENGTH: 336
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(336)
;   OTHER INFORMATION: n = A,T,C or G
US-09-854-133-560

  Query Match             10.0%;  Score 319;  DB 4;  Length 336;
  Best Local Similarity   97.9%;  Pred. No. 6.1e-57;
  Matches  330;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       2429 CTCTGCAAAAATAAAAATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTAT 2488
             |||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||
Db          1 CTCTGCAAAAATAANNATAAAAAAATAAATAAAATTTTAAAAATAAT-AAAATTCACTAT 59

Qy       2489 ATACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATC 2548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATC 119

Qy       2549 CATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTA 2608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTA 179

Qy       2609 ATGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCC 2668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCC 239

Qy       2669 CAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTG 2728
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        240 CAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGNGAAATGCTG 299

Qy       2729 TCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATG 2765
                 |||| |||||||||||||||| ||||||||||||
Db        300 NCCAGNTTTATTTTTTTTATGTTGNTATCCTTGGATG 336


RESULT 11
US-09-313-300-9
; Sequence 9, Application US/09313300
; Patent No. 6222027
; GENERAL INFORMATION:
;  APPLICANT: Kaser, Matthew, R.
;  APPLICANT: Lal, Preeti
;  APPLICANT: Yue, Henry
;  APPLICANT: Tang, Tom, Y.
;  APPLICANT: Baughn, Mariah, R.
;  APPLICANT: Azimzai, Yalda
;  TITLE OF INVENTION: MOLECULES EXPRESSED IN HIPPOCAMPUS
;  FILE REFERENCE: PB-0012 US
;  CURRENT APPLICATION NUMBER: US/09/313,300
;  CURRENT FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PERL Program
; SEQ ID NO 9
;   LENGTH: 1045
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (871)...(899)
;   OTHER INFORMATION: a or g or c or t, unknown, or other
;   FEATURE:
;   NAME/KEY:
;   OTHER INFORMATION: 700122146
;  PUBLICATION INFORMATION:
US-09-313-300-9

  Query Match              6.2%;  Score 199;  DB 3;  Length 1045;
  Best Local Similarity   63.4%;  Pred. No. 9.1e-32;
  Matches  476;  Conservative    0;  Mismatches  220;  Indels   55;  Gaps    9;

Qy       2462 ATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCAG 2521
              | |  || || ||  |||| |||||||||||||| |||    || | ||| |||| ||||
Db          1 AATAAAATAAAAAATAAAAATCACTATATACACACATA---TAAAAGAAAAAAGTGTCAG 57

Qy       2522 TTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCAA 2581
              ||||||||| ||||| |||||||||  |  |||||||||| || ||| ||||| | ||||
Db         58 TTGCAGCTACTTGTCGAAATTAATACCTGTTTCTTTTTATCTATGGTAAATATCGTGCAA 117
```

```
Qy       2582 TTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGTT 2641
Db        118 TTATAGATCTGGATTTTGAACCACTTCCTGAAAGCAGCAC-CAGAGTACTCGAAGGTGCT 176

Qy       2642 GGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGG 2701
Db        177 TGTGTTCTCTGCTGATTTGGCTGTTTCCAATGTTTACATTATTTAATCTTGCAAAAATGA 236

Qy       2702 TTCTGTGCACTTGGATGTGAAATGCTGTCCAG------TTTTATTTTTTTTATGTTGTTA 2755
Db        237 TCCTGTGCACTTGGATGTGACATGCTGTCTAGTCCGGTTTCATCTTTTTTTTTAATGTTG 296

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
Db        297 TTTATTTTTGGATGTACAAAAGAAAAATTGGGGGGAGGGGGTGATCTCTGTAGATACTCT 356

Qy       2816 CATCTTTAGAAGTTATC-----AGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGG 2870
Db        357 TGTACTTTGAAGTTACCGGAAATGGAACGGGTCTTAAAGCAGAAAGTAACTTTTCCAAGG 416

Qy       2871 AATGATACATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTG 2930
Db        417 AACAGATGCTTGCGAAGGCCCCCTTCCTTGTCTTATTCTCCAGAGACAACTGAAATTTAG 476

Qy       2931 CTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACA 2990
Db        477 CTTCTTTGTTGCAGCAAAGCTCTTTGCCCAGGTGAACACTGACCACCGCGGGTTTTCTAT 536

Qy       2991 CACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTT 3050
Db        537 GTCAGAAAGA----------------------AGAAGAAAACAAAAACATGCTCGAGCTT 574

Qy       3051 TTTCTAACTTCCCCTT-GCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTAT 3109
Db        575 TTTCTAACCTCCCCTTGGGGGTCTGTTGTGCGAACCCC-----TCTTTCTTCAATATCGT 629

Qy       3110 GTCAGTTTATTCTCTTTAATGGACTGTAAAAAA----------ATGTAATCACAAGAGT 3158
Db        630 GTCACTTTATTCTCTTTAATGGACTGTAACAAACAACAACAACAATGTAATCACGAGAGT 689

Qy       3159 GCCAAAT-TCTTGAAATGCCAAAAGGCTTTT 3188
Db        690 GCCAAATATCTTGAAACGCCAAAAGGCATTT 720
```

RESULT 12
US-09-513-999C-8212
; Sequence 8212, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8212
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 33
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 168
; OTHER INFORMATION: r=a or g
US-09-513-999C-8212

Query Match 4.9%; Score 157.6; DB 4; Length 407;
Best Local Similarity 97.1%; Pred. No. 2.6e-23;
Matches 168; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

```
Qy       3017 CACACGAAAAACAAAGAAAAAAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTT 3076
Db          1 CACACGAAAAACAAAGAAAAAAAATGCTTGAGNWTTTTCTAACTTCCCCTTGCAGTCTGTT 60

Qy       3077 GTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGGACTGT 3136
Db         61 GTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGGACTGT 120

Qy       3137 AAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAATGCCAAAAGGCTTTT 3188
Db        121 AAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAATGCCAAAAGRCATTT 173
```

RESULT 13
US-09-621-976-15158
; Sequence 15158, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15158
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15158

Query Match 4.8%; Score 154.2; DB 4; Length 315;
Best Local Similarity 98.2%; Pred. No. 1.2e-22;
Matches 164; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

```
Qy       1340 GAAGCGGGAGAGA-CAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAA 1398
Db          1 GAAGCGGGAGAGAWSAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAA 60

Qy       1399 CAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCC 1458
Db         61 CAATAAGGGCGACCGASACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCC 120

Qy       1459 GTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTC 1505
Db        121 GTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTC 167
```

RESULT 14
US-09-513-999C-10536
; Sequence 10536, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG

```
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 10536
;   LENGTH: 183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-10536

  Query Match             3.6%;  Score 114;  DB 4;  Length 183;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-14;
  Matches  114;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 129

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGC 2271
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGC 183


RESULT 15
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.2%;  Score 101.2;  DB 1;  Length 7218;
  Best Local Similarity   4.7%;  Pred. No. 4.6e-11;
  Matches   19;  Conservative  260;  Mismatches  123;  Indels    0;  Gaps    0;

Qy       1162 GGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGG 1221
              :::::::    ::::::: ::::: : ::::::::    ::::::::    ::::: : ::
Db       1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375

Qy       1222 GGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGAGACGCAGGGATGGAGGGGAGG 1281
              ::: :: :::::::::  :: ::  : :::::::::  :::: : ::::: ::::::::::
Db       1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315

Qy       1282 GGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGA 1341
              ::: : :  :::::  ::::: :: :::::  :::: ::::     :::::::::  ::
Db       1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255

Qy       1342 AGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAA 1401
              :: :::::::: :::: :::::: : : :   :::::::: : :::::  :  :::: ::
Db       1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195

Qy       1402 TAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCCGTC 1461
              ::::: ::  :: :    ::: ::::::::: :::::    ::  ::  :  ::  :
Db       1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135

Qy       1462 GCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGG 1521
              : :  :::::: :: :::::::  :::::::: ::::  :::  :::: :: ::::::::
Db       1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075

Qy       1522 CCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGGAGG 1563
                ::::  |. ||| |   ||  | |  ||||| |||   ||
Db       1074 RRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGG 1033


Search completed: March 28, 2005, 06:42:24
Job time : 535.129 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 28, 2005, 03:01:38 ; Search time 4165.09 Seconds
(without alignments)
4562.763 Million cell updates/sec

Title: US-10-071-645-3
Perfect score: 3189
Sequence: 1 ccctcccctcccccctccgcc.....gaaatgccaaaaggctttta 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 2979665951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

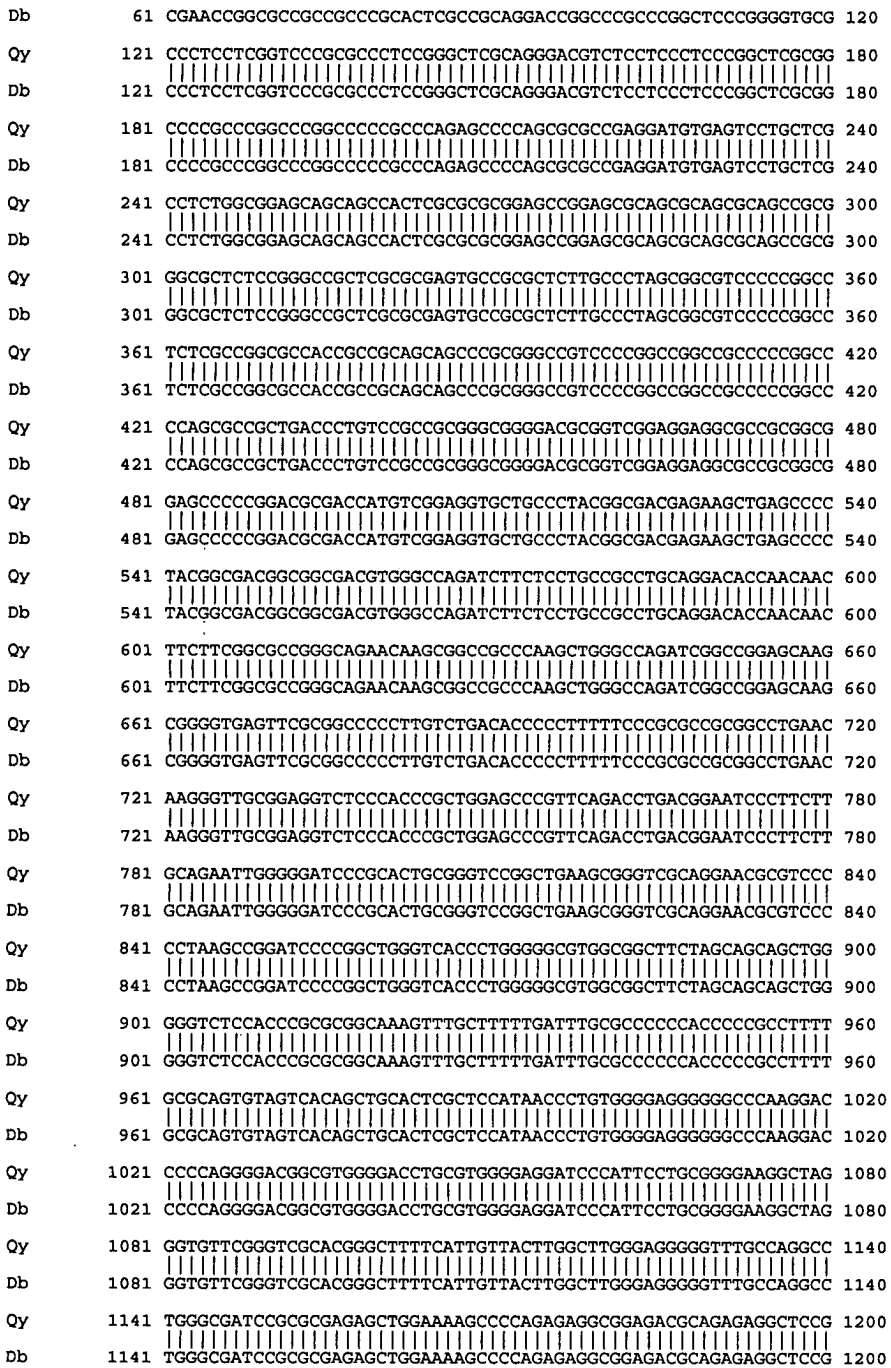
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3189	100.0	3189	15	US-10-071-645-3	Sequence 3, Appli
	2	1006.4	31.6	1850	14	US-10-198-846-11350	Sequence 11350, A
	3	1002.4	31.4	1510	10	US-09-820-649-25	Sequence 25, Appl
	4	1002.4	31.4	1510	16	US-10-160-162-25	Sequence 25, Appl
	5	1002.4	31.4	1510	19	US-10-936-773-25	Sequence 25, Appl
c	6	1002.4	31.4	2058	18	US-10-723-860-7058	Sequence 7058, Ap
	7	1000.8	31.4	1321	17	US-10-172-118-1796	Sequence 1796, Ap
	8	1000.8	31.4	1321	17	US-10-342-887-1796	Sequence 1796, Ap
	9	607.2	19.0	1104	15	US-10-071-645-1	Sequence 1, Appli
c	10	583.8	18.3	869	18	US-10-363-345A-29797	Sequence 29797, A
	11	583.8	18.3	869	18	US-10-363-345A-29798	Sequence 29798, A
	12	519	16.3	620	14	US-10-198-846-8710	Sequence 8710, Ap
	13	509	16.0	584	13	US-10-027-632-231746	Sequence 231746,
	14	509	16.0	584	17	US-10-027-632-231746	Sequence 231746,
	15	499.6	15.7	674	18	US-10-363-345A-29795	Sequence 29795, A
c	16	499.6	15.7	674	18	US-10-363-345A-29796	Sequence 29796, A
	17	488.2	15.3	869	18	US-10-363-345A-29799	Sequence 29799, A
c	18	488.2	15.3	869	18	US-10-363-345A-29800	Sequence 29800, A
c	19	414.2	13.0	549	17	US-10-152-319A-124	Sequence 124, App
c	20	413.4	13.0	417	9	US-09-736-457-475	Sequence 475, App
c	21	413.4	13.0	417	9	US-09-902-941-475	Sequence 475, App
c	22	413.4	13.0	417	9	US-09-849-626-475	Sequence 475, App
c	23	413.4	13.0	417	10	US-09-476-300-475	Sequence 475, App
c	24	413.4	13.0	417	14	US-10-017-754-475	Sequence 475, App
c	25	413.4	13.0	417	16	US-10-113-872-475	Sequence 475, App
c	26	413.4	13.0	417	17	US-10-283-017-475	Sequence 475, App
	27	399.2	12.5	417	11	US-09-978-360A-155	Sequence 155, App
c	28	394.4	12.4	453	14	US-10-060-036-2622	Sequence 2622, Ap
c	29	391.4	12.3	451	14	US-10-066-543-1967	Sequence 1967, Ap
	30	359.2	11.3	411	9	US-09-964-824A-322	Sequence 322, App
	31	359.2	11.3	411	9	US-09-954-456-313	Sequence 313, App
	32	359.2	11.3	411	9	US-09-954-456-2149	Sequence 2149, Ap
	33	359.2	11.3	411	10	US-09-873-367C-677	Sequence 677, App
	34	342.2	10.7	366	17	US-10-242-535A-49104	Sequence 49104, A
	35	342.2	10.7	366	17	US-10-085-783A-49104	Sequence 49104, A
c	36	328.2	10.3	674	18	US-10-363-345A-29793	Sequence 29793, A
	37	328.2	10.3	674	18	US-10-363-345A-29794	Sequence 29794, A
	38	319	10.0	336	9	US-09-738-973-560	Sequence 560, App
	39	319	10.0	336	9	US-09-854-133-560	Sequence 560, App
	40	319	10.0	336	15	US-10-144-649A-560	Sequence 560, App
	41	295	9.3	360	17	US-10-242-535A-44577	Sequence 44577, A
	42	295	9.3	360	17	US-10-085-783A-44577	Sequence 44577, A
	43	290.2	9.1	553	17	US-10-136-728-21	Sequence 21, Appl
	44	263	8.2	412	14	US-10-198-846-2573	Sequence 2573, Ap
	45	252.2	7.9	474	10	US-09-918-995-26636	Sequence 26636, A

ALIGNMENTS

RESULT 1
US-10-071-645-3
; Sequence 3, Application US/10071645
; Publication No. US20030148389A1
; GENERAL INFORMATION:
; APPLICANT: Bihain, Bernard
; APPLICANT: Bour, Barbara
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Schizophrenia Related Gene and Protein
; FILE REFERENCE: 92.US2.CIP
; CURRENT APPLICATION NUMBER: US/10/071,645
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/223,482
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01891
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 3
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-071-645-3

Query Match 100.0%; Score 3189; DB 15; Length 3189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTCCCCTCCCCCTCCGCCCCTCGCAGCCCCGCCGCTCGCAGCTCCCAGTCTGCCTCCC 60

Qy         61 CGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 61 CGAACCGGCGCCGCCGCCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTCCCGGGGTGCG 120
Qy 121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
Db 121 CCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCCCGGCTCGCGG 180
Qy 181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
Db 181 CCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGAGTCCTGCTCG 240
Qy 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Db 241 CCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAGCGCAGCCGCG 300
Qy 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
Db 301 GGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCGTCCCCCGGCC 360
Qy 361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
Db 361 TCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCCGCCCCCGGCC 420
Qy 421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
Db 421 CCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGGCGCCGCGGCG 480
Qy 481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
Db 481 GAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAAGCTGAGCCCC 540
Qy 541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
Db 541 TACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGACACCAACAAC 600
Qy 601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
Db 601 TTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGGCCGGAGCAAG 660
Qy 661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
Db 661 CGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCGCGGCCTGAAC 720
Qy 721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
Db 721 AAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGAATCCCTTCTT 780
Qy 781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
Db 781 GCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGGAACGCGTCCC 840
Qy 841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
Db 841 CCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAGCAGCAGCTGG 900
Qy 901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
Db 901 GGGTCTCCACCCGCGCGGCAAAGTTTGCTTTTTGATTTGCGCCCCCCACCCCCGCCTTTT 960
Qy 961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
Db 961 GCGCAGTGTAGTCACAGCTGCACTCGCTCCATAACCCTGTGGGGAGGGGGGCCCAAGGAC 1020
Qy 1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
Db 1021 CCCCAGGGGACGGCGTGGGGACCTGCGTGGGGAGGATCCCATTCCTGCGGGGAAGGCTAG 1080
Qy 1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
Db 1081 GGTGTTCGGGTCGCACGGGCTTTTCATTGTTACTTGGCTTGGGAGGGGGTTTGCCAGGCC 1140
Qy 1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
Db 1141 TGGGCGATCCGCGCGAGAGCTGGAAAAGCCCCAGAGAGGCGGAGACGCAGAGAGGCTCCG 1200
Qy 1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
Db 1201 AGAGGAGCTCCAGAGACGCGGGGACAATGAGGGGGACCGACGGCTGCAGAGAGAGACTGA 1260
Qy 1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320
Db 1261 GACGCAGGGATGGAGGGGAGGGGGTACGCTGGAGACCGAGGGTGGCAGAGACCGAGACAA 1320
Qy 1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
Db 1321 AGCTCCCGAGAGGGGAGCTGAAGCGGGAGAGACAGAGCCGAGGACGCGCGTTTGGGGAGG 1380
Qy 1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
Db 1381 ACGCAGAAGCCGCCGAAACAATAAGGGCGACCGACACCTTAGACAGGGAGAGACAGAGAC 1440
Qy 1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
Db 1441 CTCGATCGGCTGCCGGCCGTCGCGCCGAGGGACGATGGAGGGACTGAGAAAGGCGAGGCT 1500
Qy 1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
Db 1501 AAGTCGAGACGGTAAGAGAGGCCGAGGTTACGGCATGTGTCCCTGGCAGGCAGCGAAGGG 1560
Qy 1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
Db 1561 AGGCTCTGACCTCTGCGGCAGCGGGGAGCGCGGGGCGGCCGAGTCAGTCGGCCAGCGGCT 1620
Qy 1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
Db 1621 GGGAGAGGGCGCGCAGGAGGGGGCGCCCGCCCAGGCCAGGCCCTAACCCCCACCCGCTGC 1680
Qy 1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
Db 1681 GCGTCGTGGGAACCGGTTTTGGCGTCCCCTCCTGGTTCCGCTCATCTCCGCACCTAGCCT 1740
Qy 1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
Db 1741 TGCCCACCGGAGCTGCGCTCGGGACTTACCTGGGGTCCCGAGACCCAAAGACTTTGGCTC 1800
Qy 1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
Db 1801 CCTCTCCTATCCCAGCTCCAGACATTTCTGTCTAAATTAGTGCGCCTGGTGCGGGGAGGA 1860
Qy 1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
Db 1861 CGCGGGCCAGTGCGCGCCCTGGCTGCAGCAGGAGCGGCTGGGTTGGCGCCCTCTGTTTCC 1920
Qy 1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
Db 1921 TTTTCTCAGAATGGAGCTGGGACGCAGGCTGGAGGATAGAGGGTGGTGGGTGGTTCAGAG 1980
Qy 1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
Db 1981 GAAAGCAGGGAAGGGACCCCTGGCAGGGACGGAGGATGGAGCTGTTTCACCGCGCAGTGA 2040
Qy 2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
Db 2041 GCCCTGCTCCCTCGCCCTCTCCTCTCCCGACCTCCCACTCTGGGCATAACGGGAAATGTC 2100
Qy 2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
Db 2101 AGAGACCTCTGGCTAGGCCCCAGCGCGCTCACCTCTCTTTTCCCCCCTTTTTTTGCAGTT 2160
Qy 2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCTCCT 2220
Db 2161 GTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCTCCT 2220
Qy 2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280
Db 2221 GGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAACAGT 2280

```
Qy       2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCA 2340

Qy       2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGGAGG 2400

Qy       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAATAA 2460

Qy       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTCTCA 2520

Qy       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGCGCA 2580

Qy       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGGTGT 2640

Qy       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATG 2700

Qy       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTT 2760

Qy       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCT 2820

Qy       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACAT 2880

Qy       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTA 2940

Qy       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACA 3000

Qy       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTT 3060

Qy       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATT 3120

Qy       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATTCTTGAAATGCCAAA 3180

Qy       3181 AGGCTTTTA 3189
             |||||||||
Db       3181 AGGCTTTTA 3189
```

RESULT 2
US-10-198-846-11350
; Sequence 11350, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11350
; LENGTH: 1850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 26, 107, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846,
; LOCATION: 1847, 1848, 1849, 1850
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11350

Query Match 31.6%; Score 1006.4; DB 14; Length 1850;
Best Local Similarity 99.6%; Pred. No. 7.7e-214;
Matches 1030; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 326

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 386

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 446

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 506

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 566

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 626

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 686

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 746

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 806

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 866

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 926

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 986
```

```
Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 1046

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA--ACACACACAC 2995
             ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db       1047 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1106

Qy       2996 ACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCT 3055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCT 1166

Qy       3056 AACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGT 3115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGT 1226

Qy       3116 TTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAAT 3174
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1227 TTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAAT 1286

Qy       3175 GCCAAAAGGCTTTT 3188
             |||||||||| |||
Db       1287 GCCAAAAGGCATTT 1300
```

```
RESULT 3
US-09-820-649-25
; Sequence 25, Application US/09820649
; Publication No. US20030199683A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 83 Human Secreted Proteins
;  FILE REFERENCE: PZ012P1
;  CURRENT APPLICATION NUMBER: US/09/820,649
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: US/09/236,557
;  PRIOR FILING DATE: 1999-01-26
;  PRIOR APPLICATION NUMBER: PCT/US98/15949
;  PRIOR FILING DATE: 1998-07-29
;  PRIOR APPLICATION NUMBER: 60/054,212
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,209
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,234
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,218
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,214
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,236
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,215
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,211
;  PRIOR FILING DATE: 1997-07-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 1510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (39)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-820-649-25

  Query Match           31.4%;  Score 1002.4;  DB 10;  Length 1510;
  Best Local Similarity  99.2%;  Pred. No. 5.4e-213;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;
```

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db       1132 AAATGCCAAAAGGCATTT 1149
```

```
RESULT 4
US-10-160-162-25
; Sequence 25, Application US/10160162
; Publication No. US20030166541A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 83 Human Secreted Proteins
;  FILE REFERENCE: PZ012P2
;  CURRENT APPLICATION NUMBER: US/10/160,162
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: 60/295,558
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: 09/236,557
;  PRIOR FILING DATE: 1999-01-26
;  PRIOR APPLICATION NUMBER: PCT/US98/15949
;  PRIOR FILING DATE: 1998-07-29
;  PRIOR APPLICATION NUMBER: 60/054,212
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,209
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,234
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,218
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,214
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,236
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,215
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,211
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,217
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,213
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/055,968
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,969
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,972
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/056,561
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,534
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,729
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,543
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,727
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,554
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,730
;  PRIOR FILING DATE: 1997-08-19
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 1510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (39)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-160-162-25

  Query Match             31.4%;  Score 1002.4;  DB 16;  Length 1510;
  Best Local Similarity   99.2%;  Pred. No. 5.4e-213;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
                ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
                |||||||||||||| |||
Db       1132 AAATGCCAAAAGGCATTT 1149

RESULT 5
```

```
US-10-936-773-25
; Sequence 25, Application US/10936773
; Publication No. US20050037467A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 83 Human Secreted Proteins
;  FILE REFERENCE: PZ012P2
;  CURRENT APPLICATION NUMBER: US/10/936,773
;  CURRENT FILING DATE:  2004-09-09
;  PRIOR APPLICATION NUMBER: US/10/160,162
;  PRIOR FILING DATE: 2002-06-04
;  PRIOR APPLICATION NUMBER: 60/295,558
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: 09/236,557
;  PRIOR FILING DATE: 1999-01-26
;  PRIOR APPLICATION NUMBER: PCT/US98/15949
;  PRIOR FILING DATE: 1998-07-29
;  PRIOR APPLICATION NUMBER: 60/054,212
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,209
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,234
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,218
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,214
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,236
;  PRIOR FILING DATE: 1997-07-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 1510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (39)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-936-773-25

  Query Match            31.4%;  Score 1002.4;  DB 19;  Length 1510;
  Best Local Similarity  99.2%;  Pred. No. 5.4e-213;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 171

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 231

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 291

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 351

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 411

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 471

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 531

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 591

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 651

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 711

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 771

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 831

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 891

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        892 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 951

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 1011

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 1071

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1072 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1131

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db       1132 AAATGCCAAAAGGCATTT 1149


RESULT 6
US-10-723-860-7058/c
; Sequence 7058, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7058
;   LENGTH: 2058
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1952)..(1952)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (2033)..(2033)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7058

  Query Match             31.4%;  Score 1002.4;  DB 18;  Length 2058;
  Best Local Similarity   99.2%;  Pred. No. 6.3e-213;
  Matches 1030;  Conservative    0;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1739

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1679

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1619

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1559

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 1499

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 1439

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 1379

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 1319

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 1259

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 1199

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 1139

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 1079

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 1019

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       1018 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 959

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 899

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 839

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        838 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 779

Qy       3171 AAATGCCAAAAGGCTTTT 3188
              ||||||||||||| |||
Db        778 AAATGCCAAAAGGCATTT 761


RESULT 7
US-10-172-118-1796
; Sequence 1796, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1796
;   LENGTH: 1321
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: NM_018584
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1796

  Query Match             31.4%;  Score 1000.8;  DB 17;  Length 1321;
  Best Local Similarity   99.1%;  Pred. No. 1.1e-212;
  Matches 1029;  Conservative    0;  Mismatches    2;  Indels    7;  Gaps    2;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 92

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 152

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 212

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 272

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 332

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 392

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 452

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 512
```

```
Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 572

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 632

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 692

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 752

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        753 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 812

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        813 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 872

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 932

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 992

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        993 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1052

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db       1053 AAATGCCAAAAGGCATTT 1070
```

RESULT 8
US-10-342-887-1796
; Sequence 1796, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1796
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1796

Query Match 31.4%; Score 1000.8; DB 17; Length 1321;
Best Local Similarity 99.1%; Pred. No. 1.1e-212;

Matches 1029; Conservative 0; Mismatches 2; Indels 7; Gaps 2;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 92

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 152

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 212

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 272

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 332

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 392

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 452

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 512

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 572

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTTATGTTGTTATC 632

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 692

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 752

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        753 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 812

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        813 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 872

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 932

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 992

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        993 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 1052

Qy       3171 AAATGCCAAAAGGCTTTT 3188
             |||||||||||||| |||
Db       1053 AAATGCCAAAAGGCATTT 1070
```

```
RESULT 9
US-10-071-645-1
; Sequence 1, Application US/10071645
; Publication No. US20030148389A1
; GENERAL INFORMATION:
;  APPLICANT: Bihain, Bernard
;  APPLICANT:  Bour, Barbara
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: Schizophrenia Related Gene and Protein
;  FILE REFERENCE: 92.US2.CIP
;  CURRENT APPLICATION NUMBER: US/10/071,645
;  CURRENT FILING DATE:  2002-02-06
;  PRIOR APPLICATION NUMBER: 60/223,482
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: PCT/IB01/01891
;  PRIOR FILING DATE: 2001-07-26
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: Patent.pm
; SEQ ID NO 1
;   LENGTH: 1104
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 87..346
US-10-071-645-1

  Query Match             19.0%;  Score 607.2;  DB 15;  Length 1104;
  Best Local Similarity   89.2%;  Pred. No. 6.8e-125;
  Matches  787;  Conservative    0;  Mismatches   63;  Indels   32;  Gaps   11;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 311

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              |||||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db        312 -CTGGTGTCTAACT-CCCCAAAGACAATGAGTTAAGGGAGAGAATAGGAACGGCGGTAAC 369

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        370 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TAC 428

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              |||||||||||||||||||||||||||||| | ||| |||||||||| |||||| |||
Db        429 CCACGATGAAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGGCGA-AGG 486

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              |||| |||   ||| |||| |   ||||| ||||| |||| | || || |||||||| |
Db        487 AGGAAGAGAGAGAGAAGAAAGGCTTGGGC-CCTCTACAAATAAAATAAAAAAAAAAAATT 545

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
              |||||| |||||    || | |          | | ||  | |||  ||||
Db        546 TAAAATAATAAAATCCCTATA----------------TCCCATATAAGAATAAAAGAGTC 589

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
              |||||    ||  | |    ||| | | ||||||||||||||||    ||     ||
Db        590 TCAGT----GCAGTATTGGCAAAATTAAATCCATTTCTTTTTAATACGGG----AATATT 641

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
              |  |||||||||||||||||| |||||||||||||||||| || ||||||||||||||||
Db        642 GGCATTATAGATCTGGATTTTG-ACCACTTAATGAAGCGGC-ACCCCAGGTGTTTTGAGG 699

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 759

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 819

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 879

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 939

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 999

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 2997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1059

Qy       2998 ACACACACACACACACACACACACGAAAAACAAAGAAAAAAA 3039
              |||||||||||||||||||||||||||||| ||| |||||||
Db       1060 ACACACACACACACACACACACACGAAAAAAAAAAAAAAAAA 1101


RESULT 10
US-10-363-345A-29797/c
; Sequence 29797, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 29797
;   LENGTH: 869
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 29797
US-10-363-345A-29797

  Query Match             18.3%;  Score 583.8;  DB 18;  Length 869;
  Best Local Similarity   79.6%;  Pred. No. 1e-119;
  Matches  690;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy         51 TCTGCCTCCCCGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTC 110
              ||| |||||||||||||| ||||||||||||||||||||||||  |||| |||||||| |||
Db        868 TCTACCTCCCCGAACCGACGCCGCCGCCCGCACTCGCCGCAAAACCGACCCGCCCGACTC 809

Qy        111 CCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCC 170
              |||    | |||||||||||| |||||||||||||||  ||||||   ||||||||||||||||
Db        808 CCGAAATACGCCCTCCTCGATCCCGCGCCCTCCGAACTCGCAAAAACGTCTCCTCCCTCC 749

Qy        171 CGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGA 230
              || |||||| |||||||| |||| |||||||||  | |||||  |||||||   || | |
Db        748 CGACTCGCGACCCCGCCCGACCCGACCCCCGCCCAAAACCCCAACGCGCCGAAAATATAA 689

Qy        231 GTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAG 290
              ||||  |||||||||   || | || || ||||||||||||| | ||| | |||| ||||
Db        688 ATCCTACTCGCCTCTAACGAAACAACAACCACTCGCGCGCGAAACCGAAACGCAACGCAA 629

Qy        291 CGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCG 350
              |||| |||||  |||||||||  |||||||||||| | |||||||||| ||||| || ||
Db        628 CGCAACCGCGAACGCTCTCCGAACCGCTCGCGCGAATACCGCGCTCTTACCCTAACGACG 569

Qy        351 TCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCC 410
              ||||||| |||||||||  |||||||||||| || ||||||  |||||||| ||| ||
Db        568 TCCCCCGACCTCTCGCCGACGCCACCGCCGCAACAACCCGCGAACCGTCCCCGACCGACC 509

Qy        411 GCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGG 470
```

```
Db        508 GCCCCCGACCCCAACGCCGCTAACCCTATCCGCCGCGAACGAAAACGCGATCGAAAAAAA 449

Qy        471 CGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAA 530

Db        448 CGCCGCGACGAAACCCCCGAACGCGACCATATCGAAAATACTACCCTACGACGACGAAAA 389

Qy        531 GCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGA 590

Db        388 ACTAAACCCCTACGACGACGACGACGACGTAAACCAAATCTTCTCCTACCGCCTACAAAA 329

Qy        591 CACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGG 650

Db        328 CACCAACAACTTCTTCGACGCCGAACAAAACAAACGACCGCCCAAACTAAACCAAATCGA 269

Qy        651 CCGGAGCAAGCGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCG 710

Db        268 CCGAAACAAACGAAATAAATTCGCGACCCCCTTATCTAACACCCCCTTTTTCCCGCGCCG 209

Qy        711 CGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGA 770

Db        208 CGACCTAAACAAAAATTACGAAAATCTCCCACCCGCTAAAACCCGTTCAAACCTAACGAA 149

Qy        771 ATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGG 830

Db        148 ATCCCTTCTTACAAAATTAAAAAATCCCGCACTACGAATCCGACTAAAACGAATCGCAAA 89

Qy        831 AACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAG 890

Db         88 AACGCGTCCCCCTAAACCGAATCCCCGACTAAATCACCCTAAAAACGTAACGACTTCTAA 29

Qy        891 CAGCAGCTGGGGGTCTCCACCCGCGCG 917

Db         28 CAACAACTAAAAATCTCCACCCGCGCG 2


RESULT 11
US-10-363-345A-29798
; Sequence 29798, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 29798
;   LENGTH: 869
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 29798
US-10-363-345A-29798

  Query Match           18.3%;  Score 583.8;  DB 18;  Length 869;
  Best Local Similarity  79.6%;  Pred. No. 1e-119;
  Matches  690;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy         51 TCTGCCTCCCCGAACCGGCGCCGCCGCCCGCACTCGCCGCAGGACCGGCCCGCCCGGCTC 110

Db          2 TCTACCTCCCCGAACCGACGCCGCCGCCCGCACTCGCCGCAAAACCGACCCGCCCGACTC 61

Qy        111 CCGGGGTGCGCCCTCCTCGGTCCCGCGCCCTCCGGGCTCGCAGGGACGTCTCCTCCCTCC 170

Db         62 CCGAAATACGCCCTCCTCGATCCCGCGCCCTCCGAACTCGCAAAAACGTCTCCTCCCTCC 121

Qy        171 CGGCTCGCGGCCCCGCCCGGCCCGGCCCCCGCCCAGAGCCCCAGCGCGCCGAGGATGTGA 230

Db        122 CGACTCGCGACCCCGCCCGACCCGACCCCCGCCCAAAACCCCAACGCGCCGAAAATATAA 181

Qy        231 GTCCTGCTCGCCTCTGGCGGAGCAGCAGCCACTCGCGCGCGGAGCCGGAGCGCAGCGCAG 290

Db        182 ATCCTACTCGCCTCTAACGAAACAACAACCACTCGCGCGCGAAACCGAAACGCAACGCAA 241

Qy        291 CGCAGCCGCGGGCGCTCTCCGGGCCGCTCGCGCGAGTGCCGCGCTCTTGCCCTAGCGGCG 350

Db        242 CGCAACCGCGAACGCTCTCCGAACCGCTCGCGCGAATACCGCGCTCTTACCCTAACGACG 301

Qy        351 TCCCCCGGCCTCTCGCCGGCGCCACCGCCGCAGCAGCCCGCGGGCCGTCCCCGGCCGGCC 410

Db        302 TCCCCCGACCTCTCGCCGACGCCACCGCCGCAACAACCCGCGAACCGTCCCCGACCGACC 361

Qy        411 GCCCCCGGCCCCAGCGCCGCTGACCCTGTCCGCCGCGGGCGGGGACGCGGTCGGAGGAGG 470

Db        362 GCCCCCGACCCCAACGCCGCTAACCCTATCCGCCGCGAACGAAAACGCGATCGAAAAAAA 421

Qy        471 CGCCGCGGCGGAGCCCCCGGACGCGACCATGTCGGAGGTGCTGCCCTACGGCGACGAGAA 530

Db        422 CGCCGCGACGAAACCCCCGAACGCGACCATATCGAAAATACTACCCTACGACGACGAAAA 481

Qy        531 GCTGAGCCCCTACGGCGACGGCGGCGACGTGGGCCAGATCTTCTCCTGCCGCCTGCAGGA 590

Db        482 ACTAAACCCCTACGACGACGACGACGACGTAAACCAAATCTTCTCCTACCGCCTACAAAA 541

Qy        591 CACCAACAACTTCTTCGGCGCCGGGCAGAACAAGCGGCCGCCCAAGCTGGGCCAGATCGG 650

Db        542 CACCAACAACTTCTTCGACGCCGAACAAAACAAACGACCGCCCAAACTAAACCAAATCGA 601

Qy        651 CCGGAGCAAGCGGGGTGAGTTCGCGGCCCCCTTGTCTGACACCCCCTTTTTCCCGCGCCG 710

Db        602 CCGAAACAAACGAAATAAATTCGCGACCCCCTTATCTAACACCCCCTTTTTCCCGCGCCG 661

Qy        711 CGGCCTGAACAAGGGTTGCGGAGGTCTCCCACCCGCTGGAGCCCGTTCAGACCTGACGGA 770

Db        662 CGACCTAAACAAAAATTACGAAAATCTCCCACCCGCTAAAACCCGTTCAAACCTAACGAA 721

Qy        771 ATCCCTTCTTGCAGAATTGGGGGATCCCGCACTGCGGGTCCGGCTGAAGCGGGTCGCAGG 830

Db        722 ATCCCTTCTTACAAAATTAAAAAATCCCGCACTACGAATCCGACTAAAACGAATCGCAAA 781

Qy        831 AACGCGTCCCCCTAAGCCGGATCCCCGGCTGGGTCACCCTGGGGGCGTGGCGGCTTCTAG 890

Db        782 AACGCGTCCCCCTAAACCGAATCCCCGACTAAATCACCCTAAAAACGTAACGACTTCTAA 841

Qy        891 CAGCAGCTGGGGGTCTCCACCCGCGCG 917

Db        842 CAACAACTAAAAATCTCCACCCGCGCG 868


RESULT 12
US-10-198-846-8710
; Sequence 8710, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8710
;   LENGTH: 620
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 4, 5, 25, 218, 251, 275, 297, 306, 327, 340, 341, 349, 356,
;   LOCATION: 357, 381, 383, 384, 427, 437, 460, 462, 475, 482, 507, 522,
;   LOCATION: 551, 573, 575, 583, 587, 591, 608, 609, 611, 619
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8710

  Query Match            16.3%;  Score 519;  DB 14;  Length 620;
  Best Local Similarity  91.8%;  Pred. No. 2.4e-105;
  Matches  539;  Conservative    0;  Mismatches   47;  Indels    1;  Gaps    1;

Qy       2205 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 2264
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 CGACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAA 93

Qy       2265 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 2324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATT 153

Qy       2325 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 2384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 ACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAG 213

Qy       2385 ATGAGGC-GAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAA 2443
             |||| || ||||||||||||||||||||||||||||||| ||||||||||  |||||||||
Db        214 ATGANGCTGAGAGGAGGAGGAGGAGGAGGAGAAGGCTNTGGGCTCCTCTATAAAAATAAA 273

Qy       2444 AATAAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGA 2503
             | | |||||||||||||||||| ||||| || |||||||||||||||||||| ||||||
Db        274 ANTCAAAAAATAAATAAAATTTTNAAAATTATNAAAATTCACTATATACACATNTAAAGA 333

Qy       2504 AATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATAT 2563
             ||||||  ||||||| | ||||  ||||| |||||||||||||||| |  |||||||||
Db        334 AATAAANNGAAGTCTNATTTGCNNCTATTCGTCAAAATTAATATCCANTNNTTTTTATAT 393

Qy       2564 ACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACAC 2623
             |||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db        394 ACGGTGAATATTGCGCAATTATAGATCTGGATTNTGAACCACTNAATGAAGCGGCAACAC 453

Qy       2624 CAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAT 2683
             |||||| | |||||||||||| |||||| |||||||||||||||||||||||| ||||||
Db        454 CAGGTGNTNTGAGGTGTTGGCNTTCTTCNCTGATTTGGCTGTTCCCAATGTTTNCATTAT 513

Qy       2684 TTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTT 2743
             |||||||| |||||||||| |||||||||||||||| |||||||||||| | |||
Db        514 TTAATCTTNCAAAAATGGTTGTGTGCACTTGGATGTGNAATGCTGTCCAGTCTAATTGAN 573

Qy       2744 TTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTC 2790
                | ||| ||  || | | |||||||||||||||  | ||||||| |
Db        574 TNAATGATGNAATNCCTNGATGTACAAAAAATTCNNANAATGATCNC 620


RESULT 13
US-10-027-632-231746
; Sequence 231746, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 231746
;   LENGTH: 584
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(584)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-231746

  Query Match            16.0%;  Score 509;  DB 13;  Length 584;
  Best Local Similarity  98.3%;  Pred. No. 3.9e-103;
  Matches  536;  Conservative    1;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2651 CGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCA 2710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCA 60

Qy       2711 CTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAA 2770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAA 120

Qy       2771 AAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTA 2830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTA 180

Qy       2831 TCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATT 2890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATT 240

Qy       2891 TTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAG 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAG 300

Qy       2951 TTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACACACACACACACACA 3004
             |||||||||||||||||||||||||||||||||||||      ||||||||||||||||||
Db        301 TTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACACACACACACA 360

Qy       3005 CACACACACACACACACGAAAAACAAAGAAAAAAAATGCTTGAGCTTTTTCTAACTTCCCC 3064
             ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        361 CACACACACACACACACGAAAAACRAAGAAAAAAAATGCTTGAGCTTTTTCTAACTTCCCC 420

Qy       3065 TTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCT 3124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCT 480

Qy       3125 TTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAATGCCAAAAGG 3183
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        481 TTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAATGCCAAAAGG 540

Qy       3184 CTTTT 3188
                | |||
Db        541 CATTT 545


RESULT 14
US-10-027-632-231746
; Sequence 231746, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
```

```
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 231746
;   LENGTH: 584
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(584)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-231746

  Query Match          16.0%;  Score 509;  DB 17;  Length 584;
  Best Local Similarity   98.3%;  Pred. No. 3.9e-103;
  Matches  536;  Conservative    1;  Mismatches    1;  Indels    7;  Gaps    2;

Qy       2651 CGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCA 2710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCA 60

Qy       2711 CTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAA 2770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAA 120

Qy       2771 AAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTA 2830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTA 180

Qy       2831 TCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATT 2890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATT 240

Qy       2891 TTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAG 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAG 300

Qy       2951 TTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACACACACACACACACA 3004
             ||||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db        301 TTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACACACACACACACACA 360

Qy       3005 CACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTTCCCC 3064
             ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        361 CACACACACACACACACGAAAAACRAAGAAAAAAATGCTTGAGCTTTTTCTAACTTCCCC 420

Qy       3065 TTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCT 3124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCT 480

Qy       3125 TTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAATGCCAAAAGG 3183
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        481 TTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAATGCCAAAAGG 540

Qy       3184 CTTTT 3188
             | |||
Db        541 CATTT 545

RESULT 15
US-10-363-345A-29795
; Sequence 29795, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 29795
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 29795
US-10-363-345A-29795

  Query Match          15.7%;  Score 499.6;  DB 18;  Length 674;
  Best Local Similarity   83.8%;  Pred. No. 5.3e-101;
  Matches  565;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy       1051 GGAGGATCCCATTCCTGCGGGGAAGGCTAGGGTGTTCGGGTCGCACGGGCTTTTCATTGT 1110
             |||||||   |||  |||||||||||| |||||||||||||||| ||||| |||| |||||
Db          1 GGAGGATTTTATTTTTGCGGGGAAGGTTAGGGTGTTCGGGTCGTACGGGTTTTTTATTGT 60

Qy       1111 TACTTGGCTTGGGAGGGGGTTTGCCAGGCCTGGGCGATCCGCGCGAGAGCTGGAAAAGCC 1170
              || |||| |||||||||||||||  |||  |||||||| |||||||||| ||||||||
Db         61 TATTTGGTTTGGGAGGGGGTTTGTTAGGTTTGGGCGATTCGCGCGAGAGTTGGAAAAGTT 120

Qy       1171 CCAGAGAGGCGGAGACGCAGAGAGGCTCCGAGAGGAGCTCCAGAGACGCGGGGACAATGA 1230
              ||||||||||||||| ||||||| | |||||||||  |  |||||||||||||| |||||
Db        121 TTAGAGAGGCGGAGACGTAGAGAGGTTTCGAGAGGAGTTTTAGAGACGCGGGGATAATGA 180

Qy       1231 GGGGGACCGACGGCTGCAGAGAGAGACTGAGACGCAGGGATGGAGGGGAGGGGGTACGCT 1290
              |||||| |||||| || ||||||||| ||||||| |||||||||||||||||||||||| |
Db        181 GGGGGATCGACGGTTGTAGAGAGAGATTGAGACGTAGGGATGGAGGGGAGGGGGTACGTT 240

Qy       1291 GGAGACCGAGGGTGGCAGAGACCGAGACAAAGCTCCCGAGAGGGGAGCTGAAGCGGGAGA 1350
              ||||| |||||||| ||||| |||| |  |||||||||| ||||||||||||||||||||
Db        241 GGAGATCGAGGGTGGTAGAGATCGAGATAAAGTTTTCGAGAGGGGAGTTGAAGCGGGAGA 300

Qy       1351 GACAGAGCCGAGGACGCGCGTTTGGGGAGGACGCAGAAGCCGCCGAAACAATAAGGGCGA 1410
              || |||| |||||||||||||||||||||||| ||||| || ||||| ||||||||||||
Db        301 GATAGAGTCGAGGACGCGCGTTTGGGGAGGACGTAGAAGTCGTCGAAATAATAAGGGCGA 360

Qy       1411 CCGACACCTTAGACAGGGAGAGACAGAGACCTCGATCGGCTGCCGGCCGTCGCGCCGAGG 1470
              ||| | ||||| ||||||||| |||||  |||||||| || ||| ||||||| |||||
Db        361 TCGATATTTTAGATAGGGAGAGATAGAGATTTCGATCGGTTGTCGGTCGTCGCGTCGAGG 420

Qy       1471 GACGATGGAGGGACTGAGAAAGGCGAGGCTAAGTCGAGACGGTAAGAGAGGCCGAGGTTA 1530
              |||||||||||| |||||||||||||||| |||||||||||||||||||| ||||||||
Db        421 GACGATGGAGGGATTGAGAAAGGCGAGGTTAAGTCGAGACGGTAAGAGAGGTCGAGGTTA 480

Qy       1531 CGGCATGTGTCCCTGGCAGGCAGCGAAGGGAGGCTCTGACCTCTGCGGCAGCGGGGAGCG 1590
              ||| ||||||    ||| ||| |||||||||||| | |||  | ||||| ||||||||||
Db        481 CGGTATGTGTTTTTGGTAGGTAGCGAAGGGAGGTTTTGATTTTTGCGGTAGCGGGGAGCG 540

Qy       1591 CGGGGCGGCCGAGTCAGTCGGCCAGCGGCTGGGAGAGGGCGCGCAGGAGGGGGCGCCCGC 1650
              ||||||| |||| ||||||  ||||| |||||||||||||| ||||||||||  ||
Db        541 CGGGGCGGTCGAGTTAGTCGGTTAGCGGTTGGGAGAGGGCGCGTAGGAGGGGGCGTTCGT 600

Qy       1651 CCAGGCCAGGCCCTAACCCCCACCCGCTGCGCGTCGTGGGAACCGGTTTTGGCGTCCCCT 1710
              |||  |||   |||      |  || ||||||||||||||| ||||||||||     |
Db        601 TTAGGTTAGGTTTTAATTTTTATTCGTTGCGCGTCGTGGGAATCGGTTTTGGCGTTTTTT 660

Qy       1711 CCTGGTTCCGCTCA 1724
              ||||| || | |
Db        661 TTTGGTTTCGTTTA 674
```

Search completed: March 28, 2005, 12:52:58
Job time : 4169.09 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 27, 2005, 21:06:43 ; Search time 9309.24 Seconds
(without alignments)
13039.412 Million cell updates/sec

Title: US-10-071-645-3
Perfect score: 3189
Sequence: 1 ccctcccctcccccctccgcc.....gaaatgccaaaaggctttta 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1018.4	31.9	1566	3	CR604926	CR604926 full-leng
c	2	1000.8	31.4	1321	3	AF116637	AF116637 Homo sapi
	3	822.4	25.8	983	7	CN803331	CN803331 ILLUMIGEN
	4	796	25.0	976	7	CN802133	CN802133 ILLUMIGEN
	5	795.4	24.9	954	7	CN643673	CN643673 ILLUMIGEN
	6	789.6	24.8	945	7	CN803102	CN803102 ILLUMIGEN
	7	787.4	24.7	953	7	CN803224	CN803224 ILLUMIGEN
	8	785.6	24.6	954	7	CN646555	CN646555 ILLUMIGEN
	9	783.6	24.6	950	7	CN646790	CN646790 ILLUMIGEN
	10	783	24.6	952	7	CN646545	CN646545 ILLUMIGEN
	11	783	24.6	955	7	CN644007	CN644007 ILLUMIGEN
	12	782.6	24.5	954	7	CN646445	CN646445 ILLUMIGEN
	13	780.8	24.5	911	7	CN802110	CN802110 ILLUMIGEN
	14	780.6	24.5	951	7	CN647038	CN647038 ILLUMIGEN
	15	779.4	24.4	954	7	CN803392	CN803392 ILLUMIGEN
	16	777.4	24.4	1083	3	CR606015	CR606015 full-leng
	17	776.4	24.3	964	7	CN646785	CN646785 ILLUMIGEN
	18	772	24.2	953	7	CN646019	CN646019 ILLUMIGEN
	19	766.4	24.0	955	7	CN805169	CN805169 ILLUMIGEN
	20	762.6	23.9	931	5	BX419572	BX419572 BX419572
	21	757.2	23.7	955	7	CN802272	CN802272 ILLUMIGEN
	22	756.6	23.7	1020	7	CN803473	CN803473 ILLUMIGEN
	23	752.6	23.6	951	7	CN802081	CN802081 ILLUMIGEN
	24	749.8	23.5	944	7	CN803534	CN803534 ILLUMIGEN
	25	746.8	23.4	955	7	CN801954	CN801954 ILLUMIGEN
	26	745	23.4	1001	7	CN802311	CN802311 ILLUMIGEN
	27	735.2	23.1	808	5	BU570962	BU570962 AGENCOURT
c	28	728.2	22.8	747	5	BQ018543	BQ018543 UI-H-DH1-
	29	728	22.8	1026	7	CN643857	CN643857 ILLUMIGEN
	30	727.4	22.8	890	7	CO645235	CO645235 ILLUMIGEN
	31	726.6	22.8	806	7	CN642220	CN642220 ILLUMIGEN
	32	721.4	22.6	956	7	CN805265	CN805265 ILLUMIGEN
	33	721	22.6	967	7	CN641685	CN641685 ILLUMIGEN
	34	720	22.6	993	7	CN802904	CN802904 ILLUMIGEN
	35	711.4	22.3	968	7	CN805606	CN805606 ILLUMIGEN
	36	705.2	22.1	967	7	CN644185	CN644185 ILLUMIGEN
	37	704.6	22.1	997	7	CN803212	CN803212 ILLUMIGEN
	38	704.2	22.1	853	5	BU928986	BU928986 AGENCOURT
	39	699.8	21.9	914	7	CO645211	CO645211 ILLUMIGEN
	40	688.8	21.6	1004	7	CN801895	CN801895 ILLUMIGEN
	41	685.2	21.5	992	7	CN802224	CN802224 ILLUMIGEN
	42	683	21.4	695	7	CN428017	CN428017 170006000
	43	677.2	21.2	953	7	CO645289	CO645289 ILLUMIGEN
c	44	656.8	20.6	661	5	BQ000750	BQ000750 UI-H-DH1-
	45	656.8	20.6	990	7	CN803111	CN803111 ILLUMIGEN

ALIGNMENTS

RESULT 1
CR604926
LOCUS CR604926 1566 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DF038YH05 of Fetal brain of Homo sapiens (human).
ACCESSION CR604926
VERSION CR604926.1 GI:50485733
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1566)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1566)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .1566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF038YH05"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 31.9%; Score 1018.4; DB 3; Length 1566;
Best Local Similarity 99.8%; Pred. No. 1.6e-193;
Matches 1030; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 338
```

```
Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 398

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 458

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 518

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 578

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 638

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 698

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 758

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 818

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 878

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 938

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 998

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 1058

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 1118

Qy       2998 ACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAA 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 ACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAA 1178

Qy       3058 CTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTT 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 CTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTT 1238

Qy       3118 ATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTGAAATGC 3176
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1239 ATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTGAAATGC 1298

Qy       3177 CAAAAGGCTTTT 3188
             |||||||| |||
Db       1299 CAAAAGGCATTT 1310
```

RESULT 2
AF116637/c
LOCUS AF116637 1321 bp mRNA linear HTC 08-MAY-2001

DEFINITION Homo sapiens PRO1489 mRNA, complete cds.
ACCESSION AF116637
VERSION AF116637.1 GI:7959775
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1321)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhou,G., Ouyang,S., Luo,L., Bi,J., Liu,M. and He,F.
TITLE Functional prediction of the coding sequences of 121 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1321)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhou,G., Ouyang,S., Luo,L., Bi,J., Liu,M. and He,F.
TITLE Direct Submission
JOURNAL Submitted (24-DEC-1998) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing 100850, P. R. China
FEATURES Location/Qualifiers
source 1. .1321
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FLB5706"
/tissue_type="liver"
/dev_stage="fetal"
CDS complement(991. .1152)
/note="predicted protein of HQ1489"
/codon_start=1
/product="PRO1489"
/protein_id="AAF71060.1"
/db_xref="GI:7959776"
/translation="MKRESTLKFITSLLPTMKSTTCIWYQAGRQMRREEEEEEEKALG SSAKIKIKK"
ORIGIN

Query Match 31.4%; Score 1000.8; DB 3; Length 1321;
Best Local Similarity 99.1%; Pred. No. 5.2e-190;
Matches 1029; Conservative 0; Mismatches 2; Indels 7; Gaps 2;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 1230

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 1170

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 1110

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 1050

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 990

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 TAAAATTTTAAAAATAATAAAAATTCACTATATACACATATAAAGAAATAAAAAGAAGTC 930

Qy       2518 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATTGC 870

Qy       2578 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        869 GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGAGG 810

Qy       2638 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 2697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 TGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAAAA 750

Qy       2698 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 2757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTATC 690

Qy       2758 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 2817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGTCA 630

Qy       2818 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 2877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGATA 570

Qy       2878 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAG 2937
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        569 CATAGAAAAGATTTTATTTTAAAATGAGTTGTAAAGCGTGTGTTTCTTTGTTGCTGCAAG 510

Qy       2938 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAA------ACACAC 2991
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        509 CTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACACACACAC 450

Qy       2992 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 3051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCTTT 390

Qy       3052 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 3111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTCTAACTTCCCCTTGCAGTCTGTTGTGTGAGCAGCCTGTTTATTTCTCTAATATTATGT 330

Qy       3112 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAAT-TCTTG 3170
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        329 CAGTTTATTCTCTTTAATGGACTGTAAAAAAATGTAATCACAAGAGTGCCAAATATCTTG 270

Qy       3171 AAATGCCAAAAGGCTTTT 3188
              |||||||||||||| |||
Db        269 AAATGCCAAAAGGCATTT 252
```

RESULT 3
CN803331

```
LOCUS       CN803331                983 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32763 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:13297 5' similar to Bases 1 to 983 highly similar to human
            CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION   CN803331
VERSION     CN803331.1  GI:47699307
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 983)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.02. 722 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 983   Std Error: 0.00
            Plate: CL000284  row: C  column: 01
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .983
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:13297"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN
```

```
  Query Match           25.8%;  Score 822.4;  DB 7;  Length 983;
  Best Local Similarity  95.2%;  Pred. No. 3.2e-154;
  Matches  906;  Conservative    0;  Mismatches   31;  Indels   15;  Gaps    5;

Qy       2249 TTAAGGGAGAGAATAAGAACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAG 2308
              | | | ||||||||||||||||||| |||||||| |||||||||||||||||| ||||||
Db          8 TAAGGAGAGAGAATAAGAACGGCGATAACAGTTACTGGCAAAAAGCATGAAAACAGAAAG 67

Qy       2309 CACTTTGAAATTTATTACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTA 2368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CACTTTGAAATTTATTACTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTA 127

Qy       2369 TCAGGCCGGGAGACAGATGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTC 2428
              |||||||||||||||||||||||||     ||| ||||||||||||||||||||||||||||
Db        128 TCAGGCCGGGAGACAGATGAGGCG---TGAGAAGGAGGAGGAGGAGAAGGCTCTGGGCTT 184

Qy       2429 CTCTGCAAAAATAAAAATAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTAT 2488
              |||||||||||||||||| ||||||||||||||||||||||||||||||| |||||||
Db        185 CTCTGCAAAAATAAAAATAAATAAATAAATAAAATTTTAAAAATAATAAAAAATCACTAT 244

Qy       2489 AT--ACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATA 2546
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATACACACATATAAAGAAATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATA 304

Qy       2547 TCCATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACT 2606
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCCATTTCTTTTTATATACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACT 364

Qy       2607 TAATGAAGCGGCAACACCAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTT 2666
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        365 TAATGAAGCGGCAACACCAGGTATTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTT 424

Qy       2667 CCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGC 2726
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 CCCAATGTTTACATTATTTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGC 484

Qy       2727 TGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGA 2786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        485 TGTCCAGTTTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAATATGA 544

Qy       2787 TCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAA 2846
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 TCTCTGTAGATATTCTGTTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAA 604

Qy       2847 AACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGA------TTTTATTTTAAA 2900
              |||||||||||||||||||||||||||||||||||||||||      ||||||||||||
Db        605 AACAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAGATCCTTGTTTTATTTTAAA 664

Qy       2901 ATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAA 2960
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        665 ATGAGTTGTAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATTTGCCCAAGTTAATGCAAA 724

Qy       2961 TGGACACATTTTTTATGTCAGAAAAACACACACACACACACACACACACACACACACA 3020
              |||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||
```

```
Db        725 TGGACACATTTTTTATGTCAGAAACACACACACACACACACACACACACGCACACA 784

Qy       3021 CGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTCTAACTTCCCCTTGCAGTCTGTTGTGT 3080
              ||||||||||||||||||||||||||||||||||    || ||||||||||||||||||||
Db        785 CGAAAAACAAAGAAAAAAATGCTTGAGCTTTTTTCTACCTCCCCTTGCAGTCTGTTGTGT 844

Qy       3081 GAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGG--ACTGTAA 3138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        845 GAGCAGCCTGTTTATTTCTCTAATATTATGTCAGTTTATTCTCTTTAATGGGACTGTAAA 904

Qy       3139 AAAAATGTAATCACAAGAGTG--CCAAATTCTTGAAATGCCAAAAGGCTTTT 3188
              ||||||| |||||||||| ||  | ||| |||||||||| |||||||| |||
Db        905 AAAAATGAAATCACAAGAATGGCCAAAAATCTTGAAATGGCAAAAGGCCTTT 956


RESULT 4
CN802133
LOCUS       CN802133                976 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_37701 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:15099 5' similar to Bases 61 to 972 highly similar to human
            CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION   CN802133
VERSION     CN802133.1  GI:47698109
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 976)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.05.17. 751 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 976   Std Error: 0.00
            Plate: CL000204  row: F  column: 06
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .976
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:15099"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match           25.0%;  Score 796;  DB 7;  Length 976;
  Best Local Similarity   94.9%;  Pred. No. 6.3e-149;
  Matches  858;  Conservative    0;  Mismatches   35;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              || |||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
              |||||||||||||||||||||| |||||||||  |||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
              |||||||||||||      |||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
              |||||||||||| ||||| |||||||||||||||| ||||||||||||  ||  || | |
Db        845 GCTGCAAGCTATTTGCCCCAGTTAATGCAAATGGACCCATTTTTTATGTCCAAACCACCC 904

Qy       2990 ACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAATGCTTGAGCT 3049
              |||| ||| ||||||| | | |   | | | | ||||| ||| |||||||||| ||| ||
Db        905 ACACCCACCCACACACCCCCCCCGCCCCCCCCCAAAAAAAAAAAAAAAAAAATGGTTGGACT 964

Qy       3050 TTTT 3053
              ||||
Db        965 TTTT 968


RESULT 5
CN643673
LOCUS       CN643673                954 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_8340 Katze_MMBR Macaca mulatta cDNA clone IBIUW:3286
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN643673
VERSION     CN643673.1  GI:47154683
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
```

```
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 954)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.01.05. 805 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 954   Std Error: 0.00
            Plate: CL000082  row: G  column: 11
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:3286"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match           24.9%;  Score 795.4;  DB 7;  Length 954;
  Best Local Similarity   95.8%;  Pred. No. 8.3e-149;
  Matches  852;  Conservative    0;  Mismatches   26;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAACACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||||||||      ||||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             |||||||||||| ||||| ||||||||||||||||||||||||||||||||| ||||
Db        845 GCTGCAAGCTATTTGCCCCAGTTAATGCAAATGGACACATTTTTTATGTCAGAAACACAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAA 3038
             |||||||| | | ||| ||| | ||| |||   ||||| ||| ||||||
Db        905 ACACACACCCCCCCACCCACCCGCACCCACGAAAAAAAAAAAAAAAAAA 953


RESULT 6
CN803102
LOCUS       CN803102                945 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32250 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:12218 5' similar to Bases 61 to 926 highly similar to human
            CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION   CN803102
VERSION     CN803102.1  GI:47699078
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.01. 806 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 945   Std Error: 0.00
            Plate: CL000287  row: A  column: 11
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:12218"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
```

kit (catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"

ORIGIN

Query Match 24.8%; Score 789.6; DB 7; Length 945;
Best Local Similarity 95.9%; Pred. No. 1.2e-147;
Matches 845; Conservative 0; Mismatches 25; Indels 11; Gaps 3;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||| ||||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| |||||||||  ||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAANACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||| |||      ||||||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAGAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             |||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db        845 GCTGCAAGCTATTTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAACCCAC 904

Qy       2990 ACACACACACACACACACACACACACACACGAAAAACAA 3030
             |||||||||||||||| ||||| | | | |   ||| | ||
Db        905 ACACACACACACACACCCACACGCCCCCCCGAAAAACAAAA 945
```

RESULT 7
CN803224
LOCUS CN803224 953 bp mRNA linear EST 26-MAY-2004
DEFINITION ILLUMIGEN_MCQ_32542 Katze_MMBR Macaca mulatta cDNA clone
IBIUW:13763 5' similar to Bases 61 to 897 highly similar to human
CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION CN803224
VERSION CN803224.1 GI:47699200
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 953)
AUTHORS Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.04.02. 813 Q20 bases.
PCR PRimers
FORWARD: CCCTCACTAAAGGGAACAAAA
BACKWARD: CACTATAGGGCGAATTGGGTA
Insert Length: 953 Std Error: 0.00
Plate: CL000288 row: A column: 11
Seq primer: CCCTCACTAAAGGGAACAAAA
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .953
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBIUW:13763"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze_MMBR"
/note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
kit (catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"
ORIGIN

Query Match 24.7%; Score 787.4; DB 7; Length 953;
Best Local Similarity 95.3%; Pred. No. 3.3e-147;
Matches 847; Conservative 0; Mismatches 31; Indels 11; Gaps 3;

```
Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||| ||||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| |||||||||  ||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424
```

```
Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||||||||      ||||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             |||||||||||| |||||||||||||||||||||||||||||||||||| |   |||
Db        845 GCTGCAAGCTATTTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGACACCCAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAA 3038
              |   | | | | | | ||| | | | | | |||||||||| ||||||
Db        905 CCCACCCCCCCCCCCCCCCACCCCCCCCCCCCCCGAAAAACAAAAAAAAAA 953


RESULT 8
CN646555
LOCUS       CN646555                954 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_26531 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8439
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN646555
VERSION     CN646555.1  GI:47159998
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 954)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.03.10. 788 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 954   Std Error: 0.00
            Plate: CL000144  row: H  column: 08
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:8439"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.6%;  Score 785.6;  DB 7;  Length 954;
  Best Local Similarity   95.1%;  Pred. No. 7.7e-147;
  Matches  846;  Conservative    0;  Mismatches   34;  Indels   10;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db         248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db         365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGAT-----TTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTG 2930
             |||||||||||||      |||  ||||||||||||||||||||||||||||||||||||
Db         785 TACATAGAAAAGATCCTTGTTTTATTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTTG 844

Qy       2931 CTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACACA 2990
             |||||||||||| |||||||||||||||||||||||||||||||||||| |||||
Db         845 CTGCAAGCTATTTGCCCAAGTTAATGCAATTGGACACATTTTTTATGTCAGAAACACACA 904
```

```
Qy       2991 CACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAAT 3040
              ||| ||| |   | | |       | | | | ||||||||| ||||||||
Db        905 CACCCACCCCACCCCCCCACACGCCCCCCCCCAAAAAACAAAAAAAAAAAT 954


RESULT 9
CN646790
LOCUS       CN646790                950 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_27108 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8247
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN646790
VERSION     CN646790.1  GI:47160233
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.03.11. 819 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 950   Std Error: 0.00
            Plate: CL000143  row: C  column: 01
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .950
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:8247"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.6%;  Score 783.6;  DB 7;  Length 950;
  Best Local Similarity   95.9%;  Pred. No. 1.9e-146;
  Matches  850;  Conservative    0;  Mismatches   24;  Indels   12;  Gaps    4;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              || |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
              |||||||||||||||||||||||| ||||||||| |||||||||||||||||||||||||
Db         365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
              ||||||||| |||      |||||||||||||||||||||||||||||||||||||||||
Db         785 TACATAGAAGAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db         845 GCTGCAAGCTATTTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAACACAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAA 3035
              |||||||||||| | ||| ||| |||   ||  || || ||| |||
Db         905 ACACACACACAC-CCCACCCACGCACCACCAAAAACAAAAAAAAAA 949


RESULT 10
CN646545
LOCUS       CN646545                952 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_26510 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8461
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN646545
VERSION     CN646545.1  GI:47159988
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.03.09. 790 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
```

```
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 952   Std Error: 0.00
            Plate: CL000176  row: G  column: 12
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .952
                      /organism="Macaca mulatta"
                      /mol_type="mRNA"
                      /strain="Indian"
                      /db_xref="taxon:9544"
                      /clone="IBIUW:8461"
                      /sex="female"
                      /dev_stage="adult"
                      /lab_host="E. coli SOLR"
                      /clone_lib="Katze_MMBR"
                      /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                      Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                      kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                      Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.6%;  Score 783;  DB 7;  Length 952;
  Best Local Similarity   95.6%;  Pred. No. 2.6e-146;
  Matches  852;  Conservative    0;  Mismatches   25;  Indels   14;  Gaps    4;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              || |||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
              |||||||||||||||||||||| |||||||||  |||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
              |||||||||||||||||||||||||||||||||||||||||||||    | | ||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTTCTAAGGATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
              |||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
              |||||||||||| |||||||||||||||||||||| |||||||||||||||||| ||||
Db        845 GCTGCAAGCTATTTGCCCAAGTTAATGCAAATGGACCCATTTTTTATGTCAGAAACACAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAAT 3040
              ||||||||||||   | ||||| | ||| ||||||| ||| ||||||||
Db        905 CCACACACACACAC---CCCACACCCCCACCCGAAAAAAAAAAAAAAAAAT 952


RESULT 11
CN644007
LOCUS       CN644007                955 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_9278 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8814
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN644007
VERSION     CN644007.1  GI:47157450
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.01.21. 792 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 955   Std Error: 0.00
            Plate: CL000070  row: B  column: 07
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .955
                      /organism="Macaca mulatta"
                      /mol_type="mRNA"
                      /strain="Indian"
                      /db_xref="taxon:9544"
                      /clone="IBIUW:8814"
                      /sex="female"
                      /dev_stage="adult"
                      /lab_host="E. coli SOLR"
                      /clone_lib="Katze_MMBR"
                      /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                      Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                      kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                      Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.6%;  Score 783;  DB 7;  Length 955;
  Best Local Similarity   94.8%;  Pred. No. 2.6e-146;
  Matches  845;  Conservative    0;  Mismatches   35;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
```

```
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| |||||||||  ||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||||||||      |||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             |||||||  ||| ||||| |||||||||||||||||||||||| | |   | ||||||
Db        845 GCTGCAACTTATTTGCCCCAGTTAATGCAAATGGACACATTTTTTTTATGTCAGAAACAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAAAT 3040
              ||| | | | | ||| | | | | | ||| | ||||||||| ||||||||
Db        905 CCACCCCCCCCCCCCACCCCCCCCCCCCCCCACCCAAAAAACAAAAAAAAAAAT 955


RESULT 12
CN646445
LOCUS       CN646445                954 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_26221 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8553
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
            (Hs.197922), mRNA sequence.
ACCESSION   CN646445
VERSION     CN646445.1  GI:47159888
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 954)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)

COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.03.09. 788 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 954   Std Error: 0.00
            Plate: CL000118  row: D  column: 02
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:8553"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match           24.5%;  Score 782.6;  DB 7;  Length 954;
  Best Local Similarity   94.9%;  Pred. No. 3.1e-146;
  Matches  844;  Conservative    0;  Mismatches   34;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||||| |||||||||  ||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             |||||||||||| |||||||||||||| |||||||||||||||||||  ||    |||
Db        845 GCTGCAAGCTATTTGCCCAAGTTAATGCCAATGGACACATTTTTTATGTCGAAACCCCAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAA 3038
             |||| | | ||| ||| | | | | | | |    ||||  ||| ||||||
Db        905 ACACCCCCCCACCCACCCCCCCGCCCCCCCCAAAAAACAAAAAAAAAAA 953


RESULT 13
CN802110
LOCUS       CN802110                911 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_37652 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:14438 5' similar to Bases 12 to 908 highly similar to human
            CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION   CN802110
VERSION     CN802110.1  GI:47698086
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.05.17. 713 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 911   Std Error: 0.00
            Plate: CL000204  row: C  column: 03
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:14438"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.5%;  Score 780.8;  DB 7;  Length 911;
  Best Local Similarity   95.0%;  Pred. No. 7e-146;
  Matches  863;  Conservative    0;  Mismatches   32;  Indels   13;  Gaps    5;

Qy       2146 CCTTTTTTTGCAGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACC 2205
              || |||      |||||||||||||||||||||||||||||||||||||||| |||||||||
Db          6 CCCTTTAAGCGGGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAGAATATGACC 65

Qy       2206 GACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAG 2265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GACAAGGCACCTCCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAG 125

Qy       2266 AACGGCGGTAACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTA 2325
              ||||||| |||||||||| |||||||||||||||||| ||||||||||||||||||||||
Db        126 AACGGCGATAACAGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTA 185

Qy       2326 CTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGA 2385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTAGCTTGCTACCCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGA 245

Qy       2386 TGAGGCGAGAGGAGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAA 2445
              |||||||    ||| ||||||||||||||||||||||||||| ||||||||||||||||||
Db        246 TGAGGCG---TGAGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAA 302

Qy       2446 TAAAAAAATAAATAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGA 2503
              |||| |||||||||||||||||||||||||||||||||||||||| |||||||||  ||||||||||
Db        303 TAAATAAATAAATAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGA 362

Qy       2504 AATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATAT 2563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AATAAAAAGAAGTCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATAT 422

Qy       2564 ACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACAC 2623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 ACGGTGAATATTGCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACAC 482

Qy       2624 CAGGTGTTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAT 2683
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CAGGTATTTTGAGGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAT 542

Qy       2684 TTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTT 2743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TTAATCTTGCAAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTT 602

Qy       2744 TTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTG 2803
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        603 TTTATGTTGTTATCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTG 662

Qy       2804 TTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTT 2863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TTTTATTTTGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTT 722

Qy       2864 TCTAAGGAATGATACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTG 2917
              |||||||||||||||||||||||| |      |||||||||||||||||||||||||||||
Db        723 TCTAAGGAATGATACATAGAAAAAATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTG 782

Qy       2918 TGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATG 2977
              |||||||||||||||||||||||| ||||||| |||||||||||||||| |||||||||
Db        783 TGTTTCTTTGTTGCTGCAAGCTATTTGCCCAAATTAATGCAAATGGACACTTTTTTTATG 842

Qy       2978 TCAGAAAAACACACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAA 3037
              ||||||| || ||||||| | ||| ||||||| | ||||| ||  ||||| ||| |||||
Db        843 TCAGAAACAC-CACACACCCCCACCCACACAC-CCCACACCCAAAAAAAAAAAAAAAAAA 900

Qy       3038 AATGCTTG 3045
              ||||||||
Db        901 AATGCTTG 908


RESULT 14
CN647038
LOCUS       CN647038                951 bp    mRNA    linear   EST 13-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_27621 Katze_MMBR Macaca mulatta cDNA clone IBIUW:8006
            5' similar to Bases 1 to 360 highly similar to human CaMKIINalpha
```

```
            (Hs.197922), mRNA sequence.
ACCESSION   CN647038
VERSION     CN647038.1  GI:47160481
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.03.19. 778 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 951   Std Error: 0.00
            Plate: CL000175  row: F  column: 10
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:8006"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.5%;  Score 780.6;  DB 7;  Length 951;
  Best Local Similarity   94.9%;  Pred. No. 7.7e-146;
  Matches  842;  Conservative    0;  Mismatches   34;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
             ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db         248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
             || |||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db         305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
             |||||||||||||||||||| |||||||||  |||||||||||||||||||||||||||||
Db         365 TAAAATTTTAAAAATAATAAAAAATCACTATATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
             |||||||||||||      ||||||||||||||||||||||||||||||||||||||||||
Db         785 TACATAGAAAAGATCCTTGTTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
             ||||| |||||| ||||| |||||||||||||| ||||||||||||||||||||  |||||
Db         845 GCTGCCAGCTATTTGCCCCAGTTAATGCAAATGAACACATTTTTTATGTCAGAAACCCAC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAA 3036
             || | | | ||||||| | | |     | ||  || || ||| ||||
Db         905 ACCCCCCCCCACACACCCCCCCCACACCCCAAAAACAAAAAAAAAAA 951


RESULT 15
CN803392
LOCUS       CN803392                954 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32916 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:14140 5' similar to Bases 61 to 924 highly similar to human
            CaMKIINalpha (Hs.197922), mRNA sequence.
ACCESSION   CN803392
VERSION     CN803392.1  GI:47699368
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 954)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.03. 803 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 954   Std Error: 0.00
            Plate: CL000275  row: C  column: 01
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
```

```
                     /db_xref="taxon:9544"
                     /clone="IBIUW:14140"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             24.4%;  Score 779.4;  DB 7;  Length 954;
  Best Local Similarity   94.7%;  Pred. No. 1.3e-145;
  Matches  842;  Conservative    0;  Mismatches   36;  Indels   11;  Gaps    3;

Qy       2158 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 2217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGGCACCT 127

Qy       2218 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGGTAAC 2277
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        128 CCTGGTGTCTAACTCCCCCAAAGACAATGAGTTAAGGGAGAGAATAAGAACGGCGATAAC 187

Qy       2278 AGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 2337
              ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        188 AGTTACTGGCAAAAAGCATGAAAACAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC 247

Qy       2338 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCGAGAGG 2397
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        248 CCACGATGAAATCAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGGCG---TG 304

Qy       2398 AGGAGGAGGAGGAGGAGAAGGCTCTGGGCTCCTCTGCAAAAATAAAAATAAAAAAATAAA 2457
              || |||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db        305 AGAAGGAGGAGGAGGAGAAGGCTCTGGGCTTCTCTGCAAAAATAAAAATAAATAAATAAA 364

Qy       2458 TAAAATTTTAAAAATAATAAAAATTCACTATAT--ACACATATAAAGAAATAAAAAGAAG 2515
              |||||||||||||||||||||| |||||||||  ||||||||||||||||||||||||||
Db        365 TAAAATTTTAAAAATAATAAAAAATCACTATATACACACATATAAAGAAATAAAAAGAAG 424

Qy       2516 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TCTCAGTTGCAGCTATTTGTCAAAATTAATATCCATTTCTTTTTATATACGGTGAATATT 484

Qy       2576 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTGTTTTGA 2635
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        485 GCGCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCAGGTATTTTGA 544

Qy       2636 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GGTGTTGGCATTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTATTTAATCTTGCAA 604

Qy       2696 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 2755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AAATGGTTCTGTGCACTTGGATGTGAAATGCTGTCCAGTTTTATTTTTTTTATGTTGTTA 664

Qy       2756 TCCTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTTTGGT 2815
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        665 TCCTTGGATGTACAAAAAATTCAGAATATGATCTCTGTAGATATTCTGTTTTATTTTGGT 724

Qy       2816 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 2875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CATCTTTAGAAGTTATCAGGAATGTGTTTAAAACAAGAAGAGAACTTTTCTAAGGAATGA 784

Qy       2876 TACATAGAAAAGA------TTTTATTTTAAAATGAGTTGTAAAGCTTGTGTTTCTTTGTT 2929
              |||||||||||||      |||||||||| ||||||||||||||||||||||||||||||
Db        785 TACATAGAAAAGATCCTTGTTTTATTTTAATATGAGTTGTAAAGCTTGTGTTTCTTTGTT 844

Qy       2930 GCTGCAAGCTATCTGCCCAAGTTAATGCAAATGGACACATTTTTTATGTCAGAAAAACAC 2989
              |||||||||||| || || |||||||||||||||||| |||||||||||| ||| || |
Db        845 GCTGCAAGCTATTTGGCCCAGTTAATGCAAATGGACACTTTTTTTATGTCAAAAACACCC 904

Qy       2990 ACACACACACACACACACACACACACACACACGAAAAACAAAGAAAAAA 3038
                  ||| ||| ||| | | ||| |   | | |   |||| ||| |||||||
Db         905 CCACCCACCCACCCCCCCCACCCCGCCCCCCCGAAAAACAAAAAAAAAAA 953


Search completed: March 28, 2005, 06:30:21
Job time : 9313.24 secs
```